École des Hautes Études en Sciences Sociales
Ecole doctorale de l'EHESS

Université Catholique de Louvain
Commission Doctorale du Domaine Information et Communication

Thèse de Doctorat préparée dans le cadre d'une cotutelle entre l'Ecole des hautes études en sciences sociales et l'Université Catholique de Louvain

Doctoral thesis prepared as part of a joint degree between the EHESS and UCLouvain

Discipline/field **: Sciences Economiques**/Economics

Essais sur les conflits et inégalités spatiales au Sahel

Essays on conflict and spatial inequality in the Sahel

Marion Richard

Sous la direction de/under the direction of:
Flore Gubert (LEDa DIAL- Institut de Recherche pour le Developpement)
William Parienté (Université Catholique de Louvain).

Date de soutenance/date of defense **:** le 7 Juin 2024

Rapporteurs/ referees	Isabelle Chort, Université de Pau et des Pays de l'Adour Matthieu Couttenier, Ecole Normale Superieure de Lyon
Jury	Jean-François Maystadt, Université Catholique de Louvain Edward Miguel, Université de Califonie à Berkeley

Remerciements

Je souhaite profiter de la publication de cette thèse pour remercier les nombreuses personnes qui m'ont inspiré, guidé, soutenu ou tout simplement supporté (en français dans le texte) tout au long de ce parcours semé de roses et d'épines.

Bien que ce choix n'ait bien sur rien d'une expérience aléatoire, j'ai sans nul doute eu une chance inouïe de travailler sous la direction de deux incroyables chercheurs dont les connaissances n'ont d'égal que l'attention qu'ils ont porté à mes projets.

Je remercie Flore Gubert pour sa constante sollicitude, son pragmatisme imperméable à mon complicationnisme, sa disposition à confronter mes hypothèses sur la migration des femmes et les déterminants de l'éducation au Mali à la réalité du terrain, et sa générosité sans hésitation qu'il s'agisse de son temps, de ses données, de ses financements pour la recherche, ou de son infini carnet d'adresses.

Je tiens également à adresser mes sincères remerciements à William Parienté dont la rigueur analytique et la préoccupation pour les implications des résultats de la recherche ont pu je l'espère imprégner mes travaux. Son investissement dans mes projets et ma carrière de chercheuse, ses appels quotidiens pendant mes terrains au Mali pour m'aider à revoir mes questionnaires et ses incitations à être toujours plus ambitieuse m'ont permis de gagner en assurance et de trouver du sens dans la recherche.

Merci ensuite aux membres du jury, Ted Miguel, Isabelle Chort, Jean-François Maystadt et Mathieu Couttenier d'avoir accepté d'évaluer ma thèse. Mes projets ont déjà bénéficié énormément de leurs commentaires attentifs durant la pré-soutenance.

La contribution de Ted à la littérature en économie du développement est immense et m'aide à situer mes travaux dans une perspective plus large, et sa participation à ce jury de thèse m'honore. Son accueil à Berkeley m'a permis de rencontrer des étudiants et chercheurs inspirants, et ses conseils avisés m'ont permis d'améliorer mes travaux.

Je suis ravie d'avoir eu l'occasion de rencontrer Isabelle, dont les travaux sur la migration ont inspiré mes projets, et de bénéficier des ses commentaires pertinents. Je la remercie également pour son aide lors de nos collectes de données aux archives de Pau, et pour son invitation au séminaire de Pau qui nous a permis de présenter les premiers résultats de ces données, devant une audience bienveillante et interdisciplinaire.

J'ai eu beaucoup de chance que Jean-François rejoigne l'UCLouvain au début de ma thèse et amène de fantastiques projets de recherche sur la migration sur le devant de la scène de l'IRES. Merci pour son accessibilité et ses commentaires à la fois concrets et créatifs.

La lecture des travaux de Mathieu en économie des conflits a aussi été un des éléments déclencheur de mon agenda de recherche, et ce fut une grande joie de pouvoir compter sur ses commentaires perspicaces lors de l'atelier sur les conflits et durant ma pré-soutenance.

Je suis également très reconnaissante de ma collaboration avec Oliver Vanden Eynde qui m'a permis d'apprendre énormément du processus de recherche. Oliver a été un collaborateur et mentor toujours très impliqué et d'un enthousiasme pour la recherche contagieux.

Je souhaite exprimer toute ma gratitude à Gonzague Vannoorenberghe pour m'avoir aidée à percer les mystères des modèles de commerce sans aucune gravité, ainsi que pour sa gentillesse sans égale.

À tous les autres professeurs dont les commentaires et les idées ont été inestimables pour améliorer mon travail, merci - en particulier à Sylvie Lambert, Joseph Gomes, Catherine Guirkinger, Frédéric Docquier, Denis Cogneau, Karen Macours, Suanna Oh, François Libois, Akiko Suwa-Eisenmann, Liam Wren-Lewis, Luc Behagel and David Margolis.

Mes co-auteurs et amis Ismaël et Zhexun ont été des compagnons de thèse sans pareille. Je suis heureuse qu'un de nos nombreux projets ait pu aboutir dans cette thèse et j'espère que nous continuerons à explorer notre infini réservoir d'idées. Les notifications de nos perpétuelles conversations ne quittent jamais mon téléphone, me rappelant parfois à des échéances fatidiques, mais bien plus souvent à des idées, de nouveaux papiers, de nouvelles données, de nouveaux résultats qui décuplent ma motivation. Je ne sais pas si j'aurais pu survivre aux affres du confinement sans nos appels, merci à eux.

Merci à Elsa Gautrain pour l'enthousiasme partagé pour les méthodes d'enquêtes et les tissus indigos, je me réjouis à l'idée de continuer à collaborer sur nos projets au Mali.

Merci à Alex Armand, Yannik Schenk et toute l'équipe de NOVAFRICA pour m'avoir permis une douce transition entre la thèse et le post-doc. Je me réjouis de continuer à travailler ensemble!

Merci à l'équipe de l'Institut National de la Statistique du Mali, Arouna Sougane, Siaka Cissé, Aminata Coulibaly, Amadou Traoré, Ismaila Konate et Al Housseyni Haidara sans qui les travaux de cette thèse n'auraient pas été possibles. La tâche herculéenne que représente les collectes des données de l'INSTAT repose sur des personnes d'un talent et d'une force de travail impressionnantes.

Merci également à l'équipe du Centre des Archives du Personnel Militaire pour leur accueil, ainsi qu'à Beka Lomidze, Adil Boufarsi, Tesnim Kouati, Mor Cissé, Charlie Nze Essono and Pathé Diakhaté pour leur travail de fourmis sur ces archives.

Merci à mes collègues à l'Institut Goethe de Ramallah, Samira et Audrey, de Laterite, Dimitri et Anirudh et au centre de développement de l'OCDE, Hyeshin et Bram pour avoir semé les graines de ma passion pour la recherche en économie du développement.

Merci à Séverine et Virginie pour m'avoir guidée à travers les les étapes administratives de la thèse à l'UCLouvain avec bonne humeur.

Merci à toutes les organisations qui ont financé les travaux de cette thèse, y compris le FNRS à travers le programme EOS, l'EURPgSE, l'Institut de Recherche pour le Développement, les groupes d'Histoire Économique et de Développent de PSE, le fonds STEG, l'UCLouvain4Migration, l'Institut Convergences Migrations et le CEPREMAP.

Merci à mes collègues de Belgique. On n'a pas Beaubourg, ni la Seine, On n'a pas de lumière du jour, 6 mois dans l'année, mais Tiziano, Giuilia, Laura, Natalia, Gaia, Luigi, Rosa, Diego, Adam, Chiara, Esmeralda, Florian, Charles, Zhaniya, Mathilde, Nitin, Pablo, Rory, Elsa, Dorothée, Léo et Yannik ont illuminé mes journées!

Merci à Paola, Julia, Andréa, Charlotte, Oscar et Davi pour leur accueil chaleureux au bureau R6-01 lors de mes séjours à PSE, et à tous les autres fantastiques doctorants de PSE, en particulier Sebastián, Alvaro, Éric, Louise, Claire, Mélanie, Irène, Dave, Tom et Kelsey.

Merci à l'incroyable colocation de Monteresso et à Anna, mes séjours à Paris auraient été bien plus courts (ou bien chers et ennuyeux) sans leur hospitalité.

Merci à Ibrahim, Léa et Manon pour m'avoir guidée à travers les rues de Bamako.

Merci à celles et ceux qui ont fait de mon séjour à Berkeley une expérience inoubliable: Daniela, Abed, Mathieu, Rosanne, Léa, Sebastian, Milan, Eléonore, Nicolas, Jeremy, Miguel, Carmen, Sueda, Pema et Poornima.

Un immense merci à mes amis depuis mes débuts à PSE, Anna, Gisela, Ivan et Duncan, qui représentent la parfaite fusion entre mes aspirations de recherche et une source d'inspiration de vie pure.

Merci aux colocataires Amre, Daniela, Ale, Greg, Ines, Lete, Mimi, et à toutes et tous les autres de la rue Maes, Charles de Fulton Street, Luca et Gabriele de Janelas Verdes pour me faire sentir chez moi à Bruxelles, Berkeley et Lisbonne. Je ne sais pas si j'aurais pu finir cette thèse sans les petits plats et verres de vin déposés à côté de mon ordinateur, les sourires du matin et les chansons au son de la guitare d'Inès.

Merci à toutes et tous les autres, en particulier Marie, Florence, Ritha, Meryl, Delphine, Charlotte, Hélène, Laura, Aida, Abu, Aylin, Alex, Sacha, Anne-Louise, Alice et Laure, Ima, Veronika, Loïc et Daniel pour me rappeler chaque jour qu'il n'y a pas que la thèse et que, comme dirait Joseph, la vie est belle.

Enfin je remercie ma famille, ma grand-mère Jacqueline pour me faire répéter inlassablement mes projets de recherche et se parer de son plus grand sérieux pour me faire croire qu'elle me suit. Je dédie ce travail à mes parents Olivier et Françoise, qui m'attendaient déjà quand Françoise Richard-Tiberghien finissait sa thèse en immunologie à l'Université de Strasbourg. Merci à eux et à Tiphaine, Issam, Laure, Dominique et à tous les Richard qui ont accepté mon statut d'éternelle étudiante, et m'ont offert un soutien inconditionnel à travers mes vagues d'enthousiasme et mes moments de doute.

Résumé

Cette thèse s'appuie sur le contexte sahélien et utilise des outils d'analyse spatiale pour étudier les impacts et les déterminants des conflits civils, ainsi que les origines coloniales des inégalités en capital humain.

Le premier chapitre, *Conflit, routes périlleuses et migration au Mali*, vise à examiner comment l'insécurité le long des principales routes migratoires au Mali affecte la capacité des ménages à recourir à la migration comme stratégie de lissage des revenus et d'adaptation aux risques. En exploitant des données détaillées sur le stock de migrants économiques et l'origine des transferts monétaires reçus de migrants des neufs régions du Mali, j'estime d'abord les coûts supplémentaires de migration générés par l'insécurité avec un modèle de gravité. Dans un second temps, j'analyse l'effet agrégé de l'insécurité moyenne des routes migratoires relatives au destinations habituelles de chaque localité du Mali avant l'émergence du conflit civil en 2011. Les résultats montrent que le stock de migrants économiques est seulement faiblement corrélé avec l'insécurité moyenne des routes pertinentes pour les destinations de migration pré-conflit, suggérant que la réduction de l'accessibilité de certaines destinations en raison du conflit est compensée par la migration vers des destinations alternatives. Les nouveaux départs saisonniers en migration sont en revanche significativement affectés par l'insécurité sur les routes migratoires. Je constate que les ménages les plus pauvres réduisent leur réponse migratoires aux sécheresses à mesure que le nombre moyen de conflits sur les routes migratoires augmente, tandis que les ménages les plus riches réduisent leur mobilité en réponse à l'insécurité sur ces routes, quelles que soient les conditions climatiques. Je présente des preuves suggestives selon lesquelles cette réduction de la mobilité des membres du ménage se traduit par une réduction de la consommation par membre du ménage pendant la saison pré-récolte et un risque accru d'insécurité alimentaire dans les localités simultanément touchées par des épisodes de sécheresse au cours de la saison agricole précédente.

Le deuxième chapitre, *Coopération Entre Armées Nationales : Le Cas des Frontières du Sahel*, étudie comment la coopération militaire entre les armées nationales peut contribuer à la réduction de l'intensité des conflits dans les zones frontalières poreuses. Dans le contexte du Sahel, nous examinons comment la création d'une force armée internationale capable d'opérer à travers les frontières du Sahel (la Force Conjointe du G5-Sahel) a affecté la dynamique des conflits. En nous appuyant sur une analyse de régression par discontinuité, nous constatons que la mission du G5 a réduit l'intensité des conflits localement dans sa zone d'opération, surtout le long des segments frontaliers plus poreux en raison de leurs caractéristiques géographiques ou de leur composition ethnique. Une analyse supplémentaire des schémas géographiques de propagation des conflits indique que la force du G5-Sahel a facilité les opérations de sécurité dans les zones frontalières.

Le troisième chapitre, *Soldats et Travailleurs : Héritages du Service Militaire Colonial et du Travail Forcé au Mali*, étudie la conscription militaire coercitive et le régime de travail forcé associé institués par les autorités coloniales françaises dans le Mali colonial entre 1927 et 1950. En raison de la pénurie de main-d'œuvre endémique, les autorités coloniales ont eu recours à l'activation des réservistes militaires comme travailleurs forcés pour la construction d'infrastructures publiques. En numérisant plus de 180 000 fiches de conscription individuelles écrites à la main ainsi que des tableaux de conscription de l'époque coloniale, nous constatons que, historiquement, l'activation du travail forcé militaire était fortement corrélée avec la densité de population locale et les demandes spécifiques

des différents projets d'infrastructures publiques dans le Mali colonial. Pour étudier l'effet à long terme de cette conscription militaire coercitive, nous exploitons un système de loterie de l'époque coloniale où les conscrits étaient assignés de manière aléatoire au service militaire régulier versus au réserviste militaire. Nous utilisons une stratégie empirique de différences dans les différences pour analyser les effets différentiels de cette loterie en fonction des taux d'activation du travail forcé au niveau du district. Nous constatons qu'une plus grande exposition au service militaire et au travail forcé a conduit à des taux de scolarisation significativement plus faible à long terme. Cela suggère une transmission intergénérationnelle potentielle du rejet envers les institutions éducatives associées au colonialisme français.

Summary

This dissertation draws from the Sahelian context and uses tools of spatial analysis to study the impacts and determinants of civil conflict, as well as historical roots of inequalities in human capital and development outcomes.

The first chapter, *Conflict, Route Insecurity and Migration in Mali*, aims to examine how insecurity along main migration routes in Mali affects households' ability to rely on migration as a risk-coping and income-smoothing strategy. Exploiting detailed data on the stock of economic migrants and the origin of remittances received from all nine regions of Mali, I first estimate the additional migration costs generated by insecurity with a gravity model. In a second step, I analyse the aggregate effect of average route insecurity faced by each of the locality on its migration market. The findings show that the stock of economic migrants is only weakly correlated with average route insecurity relevant to pre-conflict migration destinations, suggesting that reduction in accessibility of some destinations due to conflict is offset by migration to alternative destinations. New seasonal departures in migration instead are significantly affected by insecurity on migration routes. I find that the poorest households reduce their migration response to droughts as the average number of conflicts on migration routes increases, while the richest households reduce their mobility in response to route insecurity irrespective of climatic conditions. I show suggestive evidence that this reduced mobility of household members results in lower consumption per capita during the pre-harvest season, and increased risk of food insecurity in localities simultaneously affected by episodes of droughts during the past growing season.

The second chapter, *Cooperation between National Armies: Evidence from the Sahel borders*, studies how military cooperation between national armies can contribute to the reduction of conflict intensity in porous border areas. In the context of the Sahel, we investigate how the creation of an international armed force that could operate across international borders (the G5-Sahel Joint Force) affected conflict dynamics. Relying on a regression discontinuity design, we find that the G5 mission lowered the intensity of conflict locally in its zone of operation, especially along border segments more porous due to their geographical features or ethnic composition. Further analysis of geographical conflict propagation patterns indicates that the G5-Sahel force facilitated security operations in border areas.

The third chapter, *Soldiers versus Laborers: Legacies of Colonial Military Service and Forced Labor in Mali*, studies coercive military conscription and the associated forced labor regime instituted by the French colonial authorities in colonial Mali between 1927 and 1950. Due to endemic labor scarcity, the colonial authorities resorted to activating military reservists as forced laborers for public infrastructure construction. By digitizing more than 180,000 hand-written individual conscript files as well as colonial-time district conscription tables, we find that historically speaking, the activation of military forced labor was strongly correlated with local population density and specific demands of different public infrastructure projects in colonial Mali. To study the long-term effect of this coercive military recruitment, we exploit a colonial-time lottery system where conscripts were randomly assigned to regular army service versus the military reservist at the locality level. We employ a Difference-in-Differences empirical framework to analyze the differential effects of this lottery depending on district-level forced labor activation rates. We find that greater exposure to military service and forced labor has led to significantly worse human capital outcomes over the long run. This suggests a potential inter-generational transmission of rejection toward educational institutions associated with French colonialism.

Contents

II Cooperation between National Armies: Evidence from the Sahel borders

III Soldiers versus Laborers: Legacies of Colonial Military Service and Forced Labor in Mali

General Introduction

With half of global population living below the extreme poverty line found in countries classified as fragile and conflict-affected as of 2021 (World Bank 2021), the negative impacts of conflicts are disproportionately borne by the poorest. Armed conflicts entail large costs for affected populations, not only in terms of humans lives but also through the economic repercussions of destruction and instability (Blattman and Miguel, 2010; Rohner and Thoenig, 2021). This reality underscores the importance of understanding the impact of conflict and insecurity on livelihoods and pathways out of poverty.

In addition to this concern, the impact of climate change are predicted to be greater in developing countries, which tend to be located in already warmer places (IPCC, 2022), and are less resilient to climate impacts (Dell et al., 2012). While the literature has established the role of climate change as a determinant of conflict (Burke et al., 2015) through various channels, more research is needed to understand how households facing multiple shocks are able to respond to these. This is often a difficult endeavor for applied causal analysis as all the factors are interrelated.

In the first chapter of my thesis, I propose an innovative solution to this empirical challenge by taking a step back in the spatial definition of exposure to conflict, considering indirect exposure to conflict through migration networks, whereby households are affected by conflicts that may occur far from them, but in their usual migration destinations, or along the route to this destination, and hence reduce their migration opportunities.

While this is only one of the many ways through which conflict will affect households, it is an important channel as migration in developing countries represents a primary income smoothing and ex-post risk-coping strategy to address the seasonality and uncertainty of income in agriculture-based economies (Gray and Mueller, 2012). This strategy however, might not always be available in case of liquidity constraints, leading to seasonal hunger (Bryan et al., 2014) and higher wage elasticity to productivity shocks due to lower local labour supply elasticity (Jayachandran, 2006). Furthermore, destination choice may be constrained in case of shocks (Kleemans, 2015), leading to worse migration outcomes.

Building on this literature, I find evidence that route conflicts may increase liquidity constraints and reduce migration options of households.

Using a gravity model to model the migration decision across nine regions of Mali, I first document that economic migrants avoid destinations with conflicts, and destinations that involve crossing conflict-ridden areas on their ways to reach the destination. The results are indicative of strong substitution across destinations, meaning that households redirect their economic migrants to destinations that are safer and easier to reach in safety.

In the Malian context, conflicts between rural localities and pre-conflict migration destinations reduce the effect of droughts on the probability of temporary departures in the post-harvest season for poorer households, while richer households reduce migration on average when the insecurity on their migration routes increases, with no additional effect of droughts. This constrained mobility bears significant welfare implications, as evidenced by the varied migratory response to droughts, encompassing remittances from economic migrants, departures of adult household members for permanent migration, and child fostering.

The reduced form estimates show a stark reduction in consumption per capita of households in localities which faced high average insecurity on their usual migration routes shortly before the seasonal peak of migration departures, and increase in food insecurity in localities additionally hit by droughts. The reduced-form results for post-harvest departures suggest that this may be the consequence of reduced flexibility in household composition

The findings suggest that households facing multiple shocks may struggle to cope simultaneously: poorer households may have lower capacity to mitigate their exposure to conflict by avoiding risky migration routes, and when confronted with both conflict and climatic shocks, these may reinforce each other, rendering migration financially untenable.

While these findings shed light on critical spatial and temporal priorities for security provision, determining effective strategies for providing such security remains elusive. Moreover, while regional circulation of goods and people may offer economic opportunities, the movement of actors across state boundaries may also impede security provision in border areas. Borderlands often concentrate violence, as illustrated by the fact that 18.7% of recorded violent events in 2023 occurred within 50 km of an international land border, despite these areas accounting for only 5.8% of the world population. In particular, the tri-border region, where Mali, Niger, and Burkina Faso converge without distinct physical demarcation or discontinuity in ethnic composition, has emerged as the epicenter of violence in the Sahel.

The second chapter of my thesis examines how the establishment of an international armed force capable of crossing borders, known as the G5-Sahel, influenced the intensity and spatial distribution of conflict in the region's porous border areas. Our analysis indicates that the G5 mission reduced the intensity of conflict, at least locally, within its operation zone. This effects is strongest where the borders are most porous, i.e. when the same ethnic group is present on both sides of the border and the border does not align with a major river. By studying how the spatial distribution of violence responds to trigger events, we also find that the mission facilitated security operations following French major interventions in border areas.

The geographical spread of jihadist groups and the ongoing security challenges in the Sahel region put the local improvements we document in sharp perspective. In spite of these limitations, we think it is important to document that establishing zones in which national armies share security responsibilities can change conflict dynamics. Foreign military interventions is another military approach that has been implemented in the context of civil wars, alongside other tools of external intervention : mediation and foreign aid (Rohner, 2024). The efficiency of counterinsurgencies operations on the other hand, strongly hinges on their capacity to win the hearts and minds of the population (Berman et al., 2011). Among other pitfalls, the French intervention in Mali which ousting encountered widespread support from the population in 2022, failed to gather such popular support. Results from the third chapter suggest that the rejection of the former colonial power intervention was stronger in localities most intensively exposed to coercive colonial policies : military service and military forced labour.

The historical legacy of these coercive colonial policies, implemented within the framework of conscription within the French Colonial Army in colonial Mali, is at the core of the third chapter of this thesis. While the first two chapters focus on the impacts of conflicts on households welfare and potential policy solutions to conflict propagation in porous border areas, this last chapter turns to historical roots of development in the Sahel region : colonial policies and its effects on human capital accumulation.

Several studies have highlighted lasting legacies of exposure to colonial coercion on attitudes towards modern institutions such as healthcare or justice systems (Lowes and Montero, 2021; Archibong and

Obikili, 2023). Our paper contributes to this literature by documenting how negative experiences with one colonial institution, the army, can spill over to another institution, the schooling system, particularly in a setting where both institutions trace their origins back to the colonial period. Our finding suggest that localities which had a larger share of young men serving in the French Colonial Army or in forced labour camps in Mali between 1927 and 1950 still present lower school enrolment rates in 2009. Although further investigation into the mechanisms of such persistent effect is needed, the large impact on education levels, compared with the relatively low demographic weight of military conscription for active military service and forced labour (approximately 6% of young men from cohorts of 20-year-olds), suggests that the effects likely occurred through informational and normative channels rather than solely demographic shocks.

Chapter I

Conflict, Route Insecurity and Migration in Mali

1 Introduction

Armed conflicts entail large costs for affected populations, not only in terms of humans lives but also through the economic repercussions of destructions and instability (Blattman and Miguel, 2010; Rohner and Thoenig, 2021). Disproportionately, the negative impacts of conflicts are borne by the poorest, with half of global population living below the extreme poverty line found in countries classified as fragile and conflict-affected as of 2021 (World Bank 2021). This reality underscores the importance of understanding the impact of conflict and insecurity on livelihoods and pathways out of poverty. While conflict-induced population displacements have continued to receive attention in the media and scientific literature, the effect of conflict on economic migration flows is less well understood. This paper focuses on the disruptive effect of conflict on economic migration and mobility of household members, an important but risky coping mechanism in developing countries that may become more hazardous in conflict settings.

The conflict in the Sahel stands as a paramount illustration of the prolonged civil conflicts that represented the majority of global conflicts since 1946 (Rohner and Thoenig, 2021). The central Sahel region, comprising Mali, Burkina Faso, and Niger, has witnessed an steep surge in the number of violent conflicts since the Malian army was forced out of the north of the country in 2012 by a Tuareg-led rebellion. The ensuing proliferation of jihadist groups and inter-ethnic violence have claimed the lives of approximately 11,000 people and triggered large population displacements[1]. The last population census conducted in Mali prior to the crisis in 2009 revealed a population of around seven million in the center and northern regions, which are most affected by conflict. This means that millions still reside in these areas and remain vulnerable to endemic violence affecting largely civilians, who may become collateral victims of fights between armed groups or direct targets of looting and attacks. Mali, a country whose economy relies heavily on rainfed agriculture, and where seasonal migration to the capital, the mines in the southern part of the country and neigboring countries is a common strategy to prevent seasonal hunger, has been the epicenter of the increase of violence in the Sahel in the last decade. Phone surveys conducted in the country 2021 revealed that less than 13% of residents in the center and northern regions considered the road between Bamako and their locality as "rather safe", while qualitative interviews with transport companies revealed the intermittent interruption of bus lines in conflict-affected regions or a reduction of the number of stops, rendering migration unsafe or, at best, more costly.

In this context, this study aims to examine how conflict fatalities along main migration routes affects households' ability to rely on migration as a risk-coping and income-smoothing strategy. I answer this research question using Malian household data on economic migration, remittances and household members mobility. The Malian National Institute of Statistics (INSTAT), relying on a network of locally hired enumerators, maintained its quarterly data collection activities throughout most of the country[2] during the entire 2011 to 2019 period, with the exception of the year 2012 during which surveys were suspended. In total 898 rural localites with spatial coordinates have been surveyed at least twice between 2011 and 2019. This locality panel, together with geo-referenced and dated conflict events from the Armed Conflict Location and Event Data Base (ACLED), allows to test for the short-term impact of conflicts on migration patterns and remittances received.

Exploiting detailed data on the stock of migrants and receipt of remittances from a set of nine

[1]In Mali alone, the UNHCR estimates that there were 401,736 Internally Displaced People (IDP) and 158,958 Malian refugees in asylum countries as of January 2022

[2]The Kidal region has been survey in 2011 and 2019 only, while data collection activities were suspended in Gao and Tombouctou in 2013.

potential migration destinations within Mali, I construct a dyadic dataset of bilateral migration and remittances flows to estimate a gravity model of migration decisions including conflicts on a given migration corridor as bilateral migration costs in addition to the costs traditionally estimated in the literature. I find evidence that conflicts located on the shortest route between a locality and a potential destination have a large negative effect on the stock of migrants in the destination as well as the share of households reporting receiving remittances from this destination, and on the average amount received. These effects are significantly larger in localities which experienced droughts during the past growing season, and are robust to the inclusion of origin-by-year and destination-by-year fixed effects, suggesting the effect of migration route insecurity can be isolated from other origin or destination specific shocks, but also that there is substantial substitution between migration destinations.

In a second step, I construct the average migration routes conflicts relevant to each locality migration market, weighting each migration route by its pre-conflict share of migrants. The aggregate results for indicators of economic migration estimated in the gravity model show rather weak and non significant correlations between average road insecurity and aggregate economic migration flows and remittances, consistent with the gravity model evidence of substitution between migration destinations. Using high-frequency data on post-harvest temporary and permanent departures of household members, I find that average road insecurity induces reductions in post-harvest temporary departures for localities affected by droughts. Splitting the sample between households in the bottom and top half of a wealthindex created based on housing characteristics, I find that this effect is driven by households in the bottom half of the wealthindex distribution, while for households in the upper half, route insecurity significantly reduces post-harvest departures on average, with no additional effect for households in localities affected by droughts.

Furthermore, nominal food consumption expenditures are significantly lower for localities facing higher average route insecurity, although alternative channels such as trade could explain this correlation. This series of results demonstrate that the civil conflict in Mali deteriorated migration opportunities, especially for risk-coping migration, and suggest that this may have important implications for households welfare.

This paper contributes to several strands of the economics literature. The recent literature on micro-costs of conflicts has demonstrated large impact of what has been called the cost of "fear" (Rockmore, 2017; Tapsoba, 2023). Numerous studies have documented this cost of indirect exposure to conflict, that is, situations in which households are not direct victims of conflict, but undertake costly ex-ante mitigation strategies to avoid such direct exposure. Such strategies can entail sales of looteable assets (Verpoorten, 2009), reshuffling of expenditures towards consumption and away from investment (Büttner et al., 2021), abandonment of cropland located far from the households (Boudinaud and Orenstein, 2021), and adoption of agricultural practices that minimise risk instead of maximizing profits (Arias et al., 2019). This study documents a new channel through which indirect exposure to conflict can affect households welfare through reduced migration opportunities.

This paper further builds on the rather sparse economic literature on conflict and migration. Blumenstock (2012) quantifies the magnitude and timing of the short-term effect of violent conflict events on displacement using call record data in Afghanistan. In addition to these immediate conflict-induced displacements, the effect of low intensity conflict-related insecurity on permanent outmigration has been studied in the case of the Maoist ten years long conflict in India (Shrestha, 2017), and in the case of the eleven years long Nepalese civil war (Libois, 2016). In the context of the civil war in Libya, Di Maio et al. (2023) shows that irregular migrants avoid areas with high conflict intensity. The closest related study to my research is by Chort and Rupelle (2016), who analyze migration from Mexico to the US across various migration corridors with a gravity model that incorporates criminality in the origin state as a determinant of migration flows. Their findings reveal a positive effect for states situated near the border and a negative effect for states farther away, implying that the risk of attacks during migration deter migration. Building upon this hypothesis, my paper takes a step further by disentangling the effects of conflicts at origin from conflicts located specifically on the migration routes between each origin and destination by analysing preciseley geolocated conflict events and migration routes.

The disruption of regular economic migration patterns due to insecurity, however, has not received the attention its importance for households' livelihoods in developing countries should command. Temporary migration in developing countries represents an important income smoothing and ex-post risk-coping strategy to address the seasonality and uncertainty of income in agriculture-based economies (Gray and Mueller, 2012). This strategy however, might not always be available in case of liquidity constraints, leading to seasonal hunger (Bryan et al., 2014) and higher wage elasticity to productivity shocks due to lower local labour supply elasticity (Jayachandran, 2006). Furthermore, destination choice may be constrained in case of shocks (Kleemans, 2015), leading to worse migration outcomes. In this study, I contribute to the literature on barriers to migration by investigating an additional constraint to mobility: conflict-related migration costs, and study how these costs interact with liquidity constraints.

Finally, this paper contributes to the recent literature that has emphasized the positive relationship between improved transportation infrastructure and economic development, facilitating enhanced trade and migration prospects (Aggarwal, 2018; Morten and Oliveira, 2018; Donaldson, 2018; Donaldson and Hornbeck, 2016). This paper explores instead the effect of a negative shock to migration market access, demonstrating that conflicts, by increasing migration costs and thereby diminishing access to migration opportunities, have a detrimental impact on both household well-being and their resilience when confronted with climatic shocks. The contribution of this paper lies in bridging a gravity model with micro-level data, enabling an exploration of the implications for households.

2 Context

2.1 Migration in Mali

Internal and international migration play a significant role for Malian rural households as a source of income diversification. The average Malian household is composed of 9.1 members, allowing for a portfolio approach to migration. Over the period 2011-2019, more than a quarter of Malian rural households (26%) report at least one economic migrant[3] and 19% of them receive remittances (see table 3). These remittances play a substantial economic role: yearly remittances represent 38% of a trimester of household consumption[4].

Internal migration significantly outweighs international migration in driving economic mobility and financial transfers for Malian rural households. In terms of migration flows especially, with 18% of rural households reporting an economic migrant within Mali, 9% in other African countries, and an additional 3% in the rest of the world. For remittances, 12% of rural households receive remittances from Mali, 7% from other African countries, and 3% from the rest of the world. Conditional on receiving remittances however, international migrants tend to contribute slightly more

[3]An economic migrant is defined here as a migrant that reports working at destination.

[4]Including all households, they represent 7% of a trimester consumption expenditures.

to households expenditures with the average yearly transfers from migrants in Mali representing 26% of trimester consumption, 32% for regional migrants and 39% for the rest of the world. The map of regional and internal remittances flows in figure 1 illustrates the importance of internal destinations as a source of remittances. This map together with the map of internal economic migration below further contrasts the diversity of internal migration flows with rather concentrated remittances flows from three main destinations: the capital Bamako, the region of Kayes with its growing mining industry[5] and the neighboring Ivory Coast, which experienced high rate of GDP growth over the period.

Complementary data sources available for the year 2016 displayed in table 1 for a subset of economic migrants with departure dates more than six months before the time of survey (about two thirds of them) reveals that less than ten percent of economic migrants, both internal and to ECOWAS countries, are women. About 70% of economic migrants in Mali are found in urban areas, and 67% in ECOWAS countries. The share of migrants that left less than a year ago is around 39% for migrants in Mali and 23% of migrants in ECOWAS countries, assuming that all those not reported in the migrant module are excluded because their departure date is less than 6 months prior to the survey. More information on the representativeness of this complementary detailed survey is avaible in appendix section A.1.

Table 1: Summary statistics economic migrants subsample (2016, departure $\geq$ 6mo)

	Mali			ECOWAS		
	(1) Observed (departure $\geq$ 6mo)	(2) LB all	(3) UB all	(4) Observed (departure $\geq$ 6mo)	(5) LB all	(6) UB all
	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd	mean/sd
Women	0.086	0.066	0.294	0.078	0.076	0.104
	(0.281)	(0.249)	(0.456)	(0.268)	(0.265)	(0.306)
Age	29.425	27.734	30.923	30.874	30.619	31.022
	(10.054)	(9.367)	(9.254)	(12.080)	(11.996)	(11.935)
Married	0.541	0.382	0.676	0.598	0.536	0.640
	(0.499)	(0.486)	(0.468)	(0.491)	(0.500)	(0.481)
Migrated to urban area	0.704	0.543	0.771	0.670	0.651	0.680
	(0.457)	(0.498)	(0.420)	(0.471)	(0.477)	(0.467)
Departure date $\leq$ 1 year ago	0.208	0.160	0.390	0.196	0.187	0.234
	(0.406)	(0.367)	(0.488)	(0.398)	(0.391)	(0.424)
Years since departure	7.039	5.879	7.720	8.970	8.644	9.018
	(7.151)	(6.622)	(6.396)	(9.826)	(9.705)	(9.595)
Visit at least once p/year	0.455	0.351	0.579	0.270	0.263	0.291
	(0.498)	(0.478)	(0.494)	(0.445)	(0.441)	(0.455)
Observations	651	843	843	270	278	278

Notes: Descriptive statistics based on general migration module collected in 2016 for all migrants having left their household more than six months ago. The columns observed contains information from this module for migrants also reported in the economic migration module (based on household reporting migrants in the same destination in both modules). The column LB imputes statistics for migrants from the economic migration module not observed in the general migration module, probably because they left less than six months ago, imputing a lower bound to all characteristics, (eg. $women = 0$, $age = max(agemigrants)$, $married = 0$ etc..), while the UB columns imputes the upper bound value.

These economic migrants thus include recent migrants as well as longer term migrants. Sea-

[5] A recent survey condutcted by the IOM and INSTAT in the Kayes and Sikasso regions shows that 70% of international migrants work in mines.

sonal migration is a common income smoothing strategy for rural households. With agriculture contributing to 38.7% of the GDP and employing 62.6% of the population according to World Bank data, temporary departures are prevalent in approximately 20% of rural households in the months following the harvest season (EMOP 2011-2019). Although there exists regional variation, the cropping calendar follows seasonal fluctuations of rainfall, with a similar pattern across the country: a unique rainy season lasting three to six months is followed by a longer dry season. As labour demand is lower during the dry season, both temporary and permanent departures increase in the post-harvest season.

Table 2 describes the characteristics of household members leaving in the post-harvest season, with additional data on destination and reasons for migration available in a later wave of the survey collected in 2022. While women represent less than 10% of economic migrants reported in the economic migration module, they account for abouth 50% of departures in the pre-harvest season and about 45% in the post-harvest season. Household members leaving in pre and post-harvest season are also substantially younger than economic migrants, 20 years old on average. While these correspond to different definitions of migration, the large difference in characteristics between the stock of economic migrants and fpost-harvest departures of household members may also come from a reporting bias that are likely to be more accute for the stock variables relying on respondent recall. Although collected during a period marked a military coup and restricted cross-border movements due to international sanctions against Mali, data from the 2022 EMOP survey can provide some insights on the destinations and diverse reasons for migration. First, at least 80% of these departures actually qualify as migration events, as departing members leave their commune of origin; international migration however represents only 7% of temporary and 11% of permanent departures in the post-harvest quarters. Work is the first reason for temporary migration, while marriage and family reason account for most of permanent departures.

Drawing from Azam and Gubert (2005) insights, migrants are perceived as a form of insurance, providing a safety net for households in Mali. Furthermore, adjustment in household composition, through child fostering Akresh (2009) and marriage are also used as a mitigation strategy by households. Analysis carried on in this paper confirms that departures of household members in the post-harvest season are more likely for rural households affected by droughts, supporting the idea that migration is used as a coping strategy in this context. Overall, internal and regional migration, including both long-term and seasonal movements, play a crucial role in the economic livelihoods of Malians, providing income opportunities and serving as a means of risk management.

2.2 Conflict in Mali and the Sahel

Mali has experienced, since the 2012 Tuareg-led rebellion aiming for the independence of the north of the country, a general power vacuum which has not been solved by international and French interventions, and paved the way for the expansion of jihadist and ethnic self-defense groups. The number of violent events occurring on its territory has dramatically increased, and 40% of those events target civilians (ACLED). Among the violent events reported in the database, civilians can be victims of direct attacks from armed groups (reprisals, refusal to pay jihadist taxes, enforcement of sharia) or reprisal, lootings and attacks from self-defense group and government army, as well as collateral damage during clashed between government and international forces with local armed groups (e.g. IED). In total, 11394 have died since 2011, including 3683 civilians. Figure 2 displays the violent events reported in the ACLED database in Mali and neighboring countries for the years 2011 and 2019.

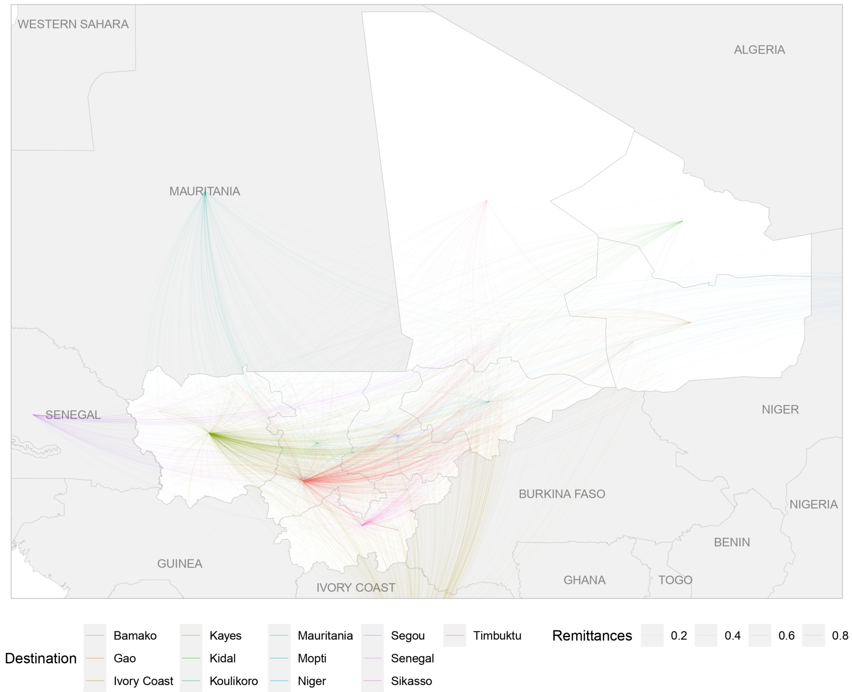
Table 2: Summary statistics households members' mobility

	Pre-harvest quarter		**Post-harvest quarters**	
	Temporary	**Permanent**	**Temporary**	**Permanent**
Sample: EMOP 2011-2019				
Women	0.497	0.501	0.448	0.451
	(0.500)	(0.500)	(0.497)	(0.498)
Age	18.933	20.168	20.701	21.852
	(14.171)	(13.918)	(14.140)	(17.693)
Married	0.283	0.301	0.296	0.373
	(0.450)	(0.459)	(0.457)	(0.484)
Observations	2957	737	7462	6135
Sample : EMOP 2022				
Women	0.407	0.671	0.444	0.627
	(0.493)	(0.473)	(0.497)	(0.486)
Age	19.883	16.370	25.008	18.297
	(13.115)	(12.009)	(19.178)	(13.323)
Married	0.198	0.132	0.303	0.110
	(0.399)	(0.340)	(0.460)	(0.313)
Left commune	0.891	0.878	0.826	0.801
	(0.313)	(0.329)	(0.380)	(0.400)
Migrated abroad	0.096	0.165	0.066	0.111
	(0.296)	(0.373)	(0.249)	(0.315)
Reason for migration				
Marriage	0.059	0.286	0.033	0.341
	(0.236)	(0.454)	(0.180)	(0.476)
Studies	0.235	0.066	0.052	0.068
	(0.425)	(0.250)	(0.221)	(0.253)
Work	0.390	0.198	0.383	0.159
	(0.489)	(0.401)	(0.486)	(0.367)
Insecurity	0.021	0.066	0.027	0.015
	(0.145)	(0.250)	(0.161)	(0.123)
Health	0.027	0.000	0.148	0.023
	(0.162)	(0.000)	(0.355)	(0.150)
Family	0.176	0.286	0.236	0.348
	(0.382)	(0.454)	(0.425)	(0.478)
Other	0.091	0.099	0.121	0.045
	(0.288)	(0.300)	(0.327)	(0.209)
Observations	187	91	601	132

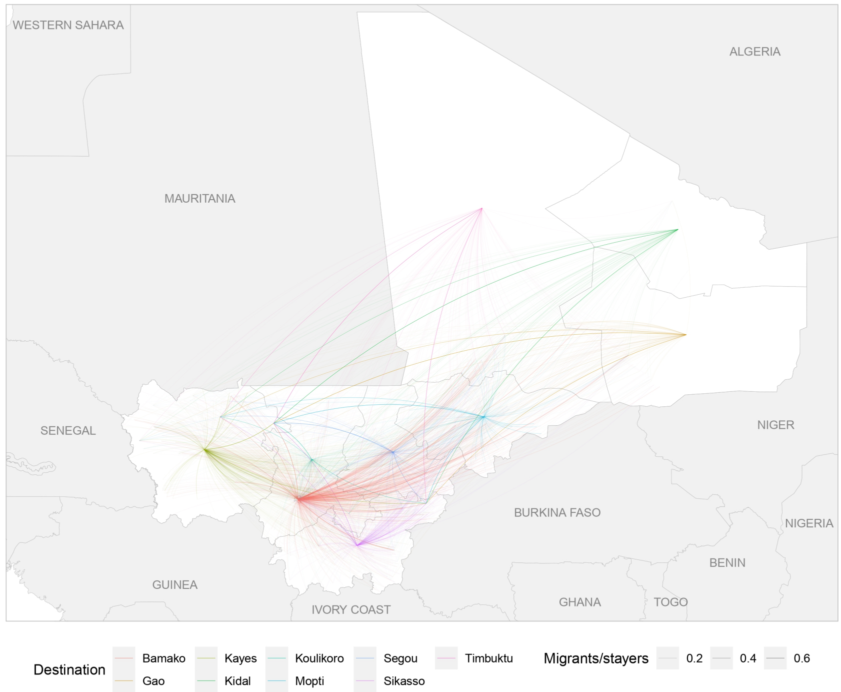
Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The p-values displayed at the bottom of the table are only for the route conflict coefficients, and conley standard errors allow fo spatial correlations up to 50km. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflict fatalities includes all fatalities within a 50 km buffer of the considered locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflict fatalities, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

Figure 1: Internal migrants stocks and regional remittances to Mali.

(a) Share of households receiving remittances (internal and main neighboring countries).



(b) Migrants per adult household members (internal migrants only).



Note: Source: EMOP data 2011-2019. Values are averaged for each origin-destination pair over the 2011-2019 period. Each flow between a rural village of Mali and an internal or regional destination is represented by a curve with color coded by destination. Destination specific data on migrant stocks are available for internal migrants only, while destination specific remittances data are available for Malian regions and Mauritania, Senegal, Ivory Coast and Niger.

According to an additional phone survey conducted in 2021 in partnership with the INSTAT on a sample of about 400 households, less than 13% of residents in the center and north regions[6], where most conflicts events occur, considered the road between Bamako and their locality as "rather safe". In addition, the 2018 Afrobarometer survey reveals that in these same regions, more than 65% of respondents have felt unsafe in their neighborhood at least once over the past year, while 50% feared for crimes happening in their homes and 28% have actually experienced theft. Qualitative interviews conducted by phone with transport companies revealed the intermittent interruption of bus lines in conflict-affected regions or a reduction of the number of stops, rendering migration unsafe or, at best, more costly.

3 Theoretical Framework

I model road insecurity as an additional migration cost with the canonical random utility maximization (RUM) model used in the migration literature (Mayda, 2010; Beine, Docquier, et al., 2011; Bertoli and Fernández-Huertas Moraga, 2013; Beine, Bertoli, et al., 2016; Beine, Parsons, et al., 2021; Ortega and Peri, 2021a) to describe an individual location decision.

The utility of an individual i migrating from origin o to destination d in the post-harvest period t where d belongs to a set of potential locations L, can be represented as a function the utility in the location $v_{d,t}$ the cost of moving to the location and an individual-specific stochastic terms ε_{idt}.

$$v_{odt} = v_{dt} - c_{odt} + \varepsilon_{idt} \tag{1}$$

Assuming that the error terms of the latent utilities $\{\varepsilon_o, \varepsilon_d, \ldots, \varepsilon_J\}$ are independent draws from a type-I extreme value distribution (MacFaden 1974), the probability that an individual i decides to migrate to a location d over all set of potential locations $L = \{o, d, \ldots, J\}$ can be written as:

$$Pr\left(y_{it} = d\right) = \frac{e^{v_{dt} - c_{odt}}}{\sum_{l \in L} e^{v_{lt} - c_{ldt}}} \tag{2}$$

This assumption implies that the odd ratio of probability of migration to any two locations, for example destination d and staying in locality o, is independent of the availability of other alternatives.

By the law of large numbers, the individual probability corresponds to the aggregate population share migrating to that particular location, which is multiplied by the total population of o to recover the total number of migrants between o and d:

$$m_{odt} = \frac{e^{v_{dt} - c_{odt}}}{\sum_{l \in L} e^{v_{lt} - c_{ldt}}} s_{ot} \tag{3}$$

So that the total number of migrants to any destination d m_{odt} is the product of the total origin population s_{ot} and the utility of migrating to this destination divided by a terms referred to in the literature as *multilateral resistance to migration*, $\sum_{l \in L} e^{v_{lt} - c_{ldt}}$, which represents the opportunity cost of migration to d.

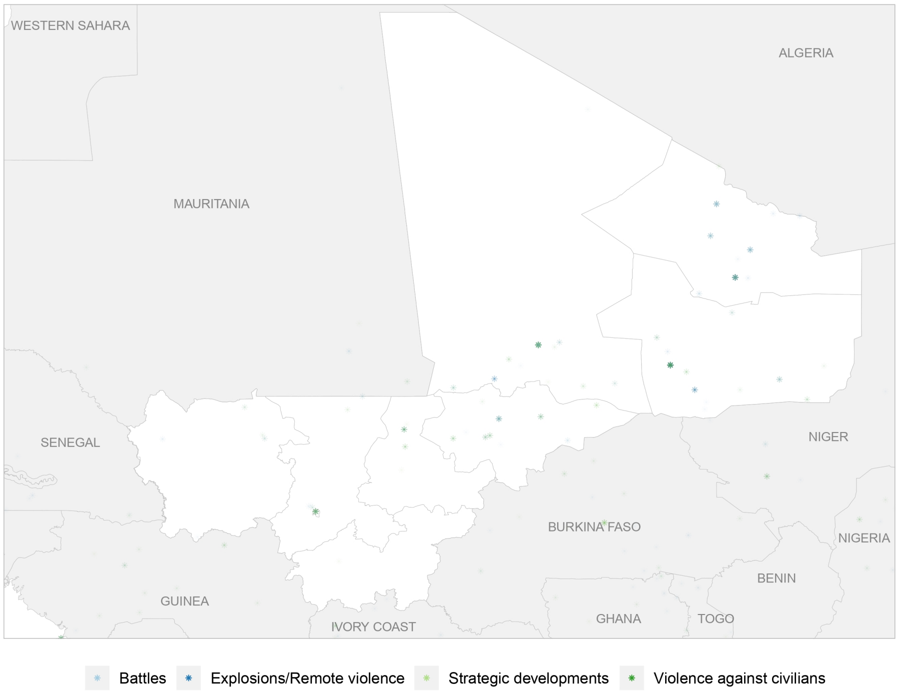
It can further be simplified by dividing the number of migrants by the number of stayers :

$$\frac{m_{odt}}{m_{oot}} = e^{v_{dt} - v_{ot} - c_{odt}} \tag{4}$$

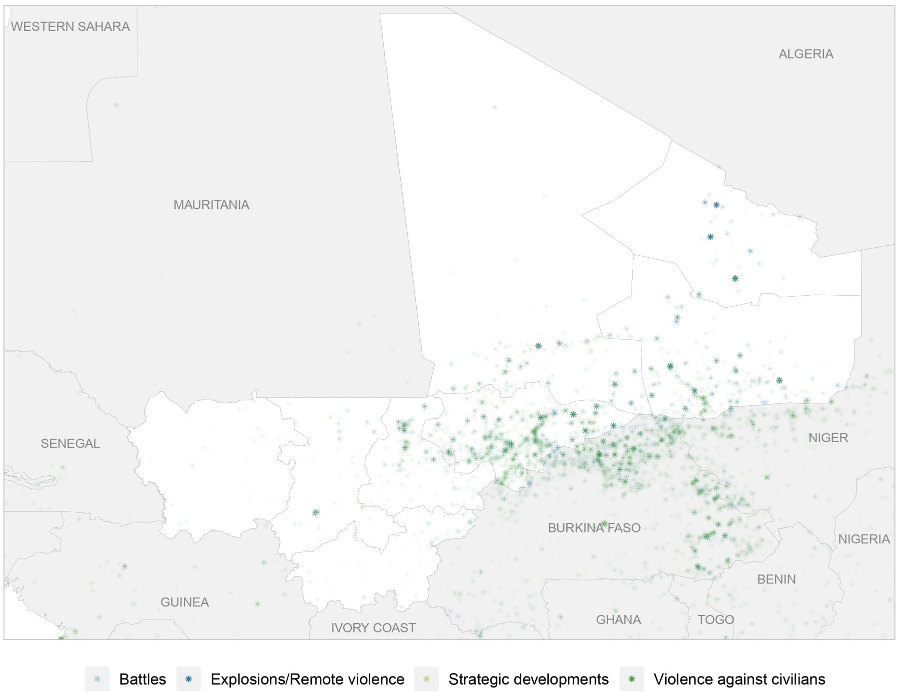
[6] Center regions: Segou and Mopti; Northern regions: Gao,Tombouctou and Kidal

Figure 2: Spatial extension of conflict during the 2011-2021 period.

(a) Conflict events 2011-2013.



(b) Conflict events 2014-2021.



Note: Source: ACLED. All the conflict events recorded in the ACLED database for the period 01/01/20 to 09/07/2021 with a transparency scale increasing in the number of fatalities.

The expected utility in any location is an increasing function of local wages $\omega_{d,t}$ and a decreasing function of local insecurity $\mu_{d,t}$, so that $v_{d,t} = f(\mu_{d,t}, \omega_{d,t})$. The expected bilateral migration costs are a increasing function of insecurity μ_{odt} on the route between o and d, distance Ω_{odt} between the two and a decreasing function of the share of individuals that speak the same language π_{od} so that $c_{odt} = f(\mu_{odt}, \Omega_{odt}, \pi_{od})$. I hypothesize that the expected level of insecurity is a function of past insecurity $E[\mu_t] = f(c_{t-1}, \epsilon_t)$, and that the agent observes realization of μ in the pre-harvest period $t-1$ and takes the decision whether to migrate in the post-harvest period t. The functional forms of these variables is made explicit in the empirical section.

4 Data

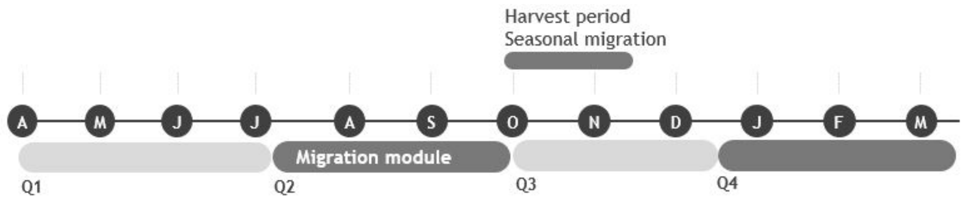
4.1 Data sources

Household survey data The "Enquête Modulaire et Permanente Auprès des Ménages" (EMOP) is collected quarterly by the Malian National Statistics Institute (INSTAT), every three months between April and January. Each wave contains information about around five thousands households, including three thousands rural households, surveyed four times over the course of a year but not followed from one year to another. The frequency of data collection allows for an objective and timely measurement of migration departures. A post-harvest departure is recorded when a household member present during the second quarter of data collection, held before the harvest, is reported as having left the household during the third or fourth quarter of data collection, around or after harvest time.

A specific module on migration and remittances conducted in the second quarter surveys the household main respondent on the current stock of migrants living and working abroad or in Mali and the amount and origin remittances received over the past 12 months. While this data is less objective than the measurement of migration based on observed departures between quarters of data collection, the migrants stocks and remittances are reported separately for each of the nine regions of Mali. This allows to create pairwise data between each locality of origin and a set of 9 potential migration destinations for internal migration. It is collected in each survey except for 2013.

To measure perception of insecurity, I use a specific module on governance, peace and security collected by Calvo et al. (2020). This module, collected in 2014, 2015 and 2016, 2017 and 2018, surveyed three adults per household. Among survey questions, information is collected on whether individuals think that there are tensions in their community, whether they think it is likely that they will become victim of violence in the future, and whether they feel unsafe in four different situations (at home and in their neighborhood by day and night). To assess the effect of route conflict on household welfare, I further exploit the consumption and food security modules of the survey. Consumption expenditure data is collected quarterly for about 20 food items and about 80 non-food items that are consistent across surveys. For each food item, expenditures in local currency over the past seven days for the whole household are reported by the household main respondent, either directly in local currency for purchased items, or in quantities valued by the enumerator given prices in local markets for self-consumption. The total value expenditures for each non food-items are directly reported for a three months period in local currency. In addition to this, the household main respondent reports whether the household has had difficulty to feed itself over a recall period of one year.

Figure 3: Data collection calender



Census data I use the last pre-conflict Malian census conducted in 2009 to observe village-level migration destinations reported by return migrants. As the focus of this study is regional and internal migration, I build village-level share of migrants accounted for by each destination as either a district within Mali or any neighboring country; together these destinations account for more than 80% of all reported past migration episodes. Moreover, leveraging the census data, I construct a variable recording the share of individuals sharing the origin main language as their mother tongue for each locality-destination pair. For a locality, the main language group is the one reported as mothertongue by the largest share of the population (on average, the first language is reported by 80% of the population) and the final variable is then the share of individuals reporting the same language as mothertongue in each destination.

Conflict data The Armed Conflict Location Event Data (ACLED) records violent events sourced from traditional media, social networks and NGOs and international organisations reports, which are geo-coded, dated and classified by type and actors involved. To focus on dynamics related to civic conflict, the analysis includes events classfied as battles, explosions and remote violence, and violence against civilians, while protests events are excluded from the sample. Only events coded with the highest geographical precision level are considered.

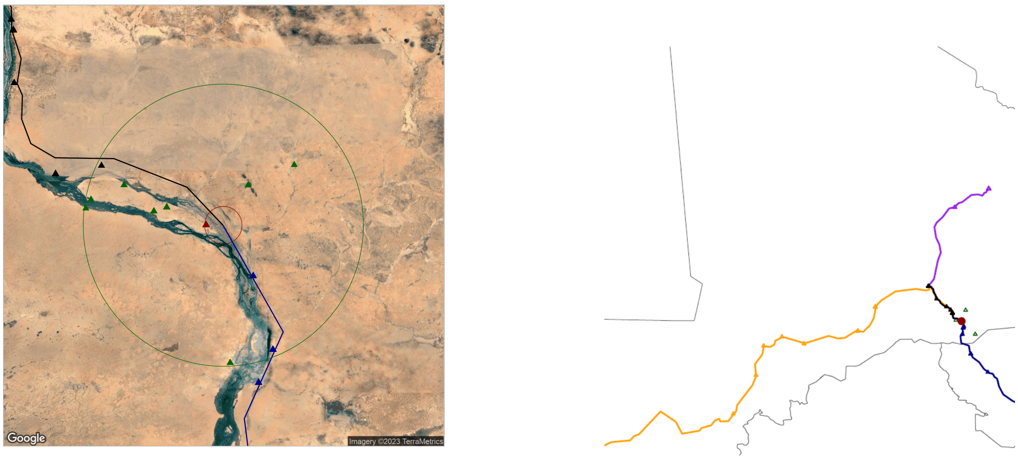
Road data The road network used to estimate the shortest route between each village and potential migration destination is based on OpenStreetMap, an open-source and crowd-sourced map of the world hosted at University College London.The base road network for Mali is the World Roads shapefile, which is part of the Vector Map Level 0 database (2000), itself based on US National Imagery and Mapping Agency's (NIMA) Operational Navigation Chart (ONC) series, and is supposed to be regularily updated by a community of volunteer mapers using satellite data, traditional maps and field observations. Using this road network, the ORSM package in R computes the shortest route between pairs of coordinates, accounting for speed limits, road quality and curvature of the road. One limitation of the road data is that most roads are unclassified, meaning that paved roads cannot be distinguised from laterites; as a result all paths are included in the analysis as potential roads, meaning that the separation of "road" conflict from non-road conflict is relative to the considered localities and destinations rather than a conflict attribute, as virtually all conflicts occur within 2km of a way in the OpenStreetMap network.

Drought data The Standardised Precipitation Evapotranspiration Index (SPEI) data from the Climatic Research Unit of the University of East Anglia combines information on precipitations and temperature with a 0.5 x 0.5 degrees precision level. With the SPEI value I construct a drought frequency variable which records the total number of dry months, for which the SPEI index is below -1.5, during the rainy season (June to November), measured at the locality level.

4.2 Construction of the two main datasets

Road, origin and destination conflicts Using the coordinates of localities included in the housheold survey, I compute for each locality and year the number of conflicts and fatalities occuring within a 15 and 25 and 50 km buffer between the first and second quarter of data collection, or alternatively, in the six months preceding the regional average date of harvest. Within the same time frames, for each locality, I recover the conflict events and fatalities occuring on the roads between the locality and several potential migration destinations: the largest city in each region of Mali, and the closest city within Côte d'Ivoire, Niger, Senegal and Mauritania. Figure 4 illustrates the classification of conflicts around a village in the region of Gao in panel (a) with a 2km radius drawn in red, a 25 km radius circle drawn in green, conflicts on the road to Niger in blue, on the road to Gao in black, and conflict on neither road in green. Panel (b) displays the same village and conflicts on the road to Niger in blue, to Gao in black, to Kidal in purple and to Bamako in orange. For each of potential destination, I compute the number of conflicts and fatalities occuring in the whole country or region for the same periods.

Figure 4: Spatial definition of road conflicts and other conflicts



(a) Spatial extent of conflicts (b) Road conflicts to several destinations

Note: Illustration of geographical definition of road conflicts for the village of Tassiga, Gao. Conflicts reported in the ACLED database for the period 2010-2023. The red (green) circle represents a 2km (25km) radius around the village. Roads between the village and Niger, Gao, Bamako and Kidal closest cities are obtained with OpenStreetMap.

Locality panel Using the eight years of survey data over the 2011 and 2014 to 2019 period, I first construct an (unbalanced) panel of 898 rural localities observed at least for two years, excluding urban areas whose livelihoods are not based on agriculture and hence seasonal migration is not practiced. Table 3 describes the main variables of interest for the localities included in the study sample for the entire period.

Dyadic panel Building on the locality panel, I construct a dyadic dataset with one observation for each locality, year, and potential destination, resulting in a dataset of 25839 observations : number of localities [898] x potential destinations [9 -(*13*)] x the number of years each locality is observed [2-8]. Table 4 describes the main variables of interest for the different samples of

Table 3: Descriptive statistics simple panel

	N	mean	sd
	Panel A: Economic migrants and remittances		
Household reports at least one economic migrant:			
Any destination	18712	0.259	0.438
Mali	18712	0.181	0.385
Other African countries	18712	0.101	0.301
Rest of the world	18712	0.028	0.164
Household receives remittances:			
Any destination	18712	0.194	0.396
Mali	18712	0.124	0.329
Other African countries	18712	0.070	0.256
Rest of the world	18712	0.027	0.163
Bamako	18712	0.069	0.254
Kayes	15556	0.035	0.185
Ivory Coast	18712	0.030	0.171
Remittances as share of trimester expenditures:			
Any destination	18712	0.075	0.268
Any destination $\mathbb{1}\{>0\}$	3888	0.387	0.500
Mali	18712	0.032	0.133
Mali $\mathbb{1}\{>0\}$	2335	0.261	0.289
Other African countries	18712	0.023	0.136
Other African countries $\mathbb{1}\{>0\}$	1488	0.328	0.403
Rest of the world	18712	0.011	0.086
Rest of the world $\mathbb{1}\{>0\}$	612	0.408	0.332
	Panel B: Post-harvest departures		
All migration, post-harvest	185217	0.051	0.219
Temporary migration, post-harvest	185217	0.037	0.189
Permanent migration, post-harvest	185217	0.013	0.114
Household attrition	19219	0.027	0.163
	Panel C: Exposure to climate shocks		
Any month of severe drought (SPEI <-1.5) growing season	18712	0.210	0.407
SPEI drought frequency growing season (severe)	18712	0.341	0.731
	Panel D: Exposure to conflict events		
Any conflict event	18712	0.210	0.407
All conflicts, 50km (count)	18712	0.424	1.287
	Panel E: Perceived insecurity		
Tensions in community	22534	0.129	0.335
Likely victim of criminality in the future	22534	0.205	0.404
Feels unsafe, count	22534	1.075	1.602
	Panel F: Other characteristics		
Household size	18712	9.059	6.365
Total household pre-harvest consumption expenditures per capita	18712	87782.303	59834.716
Total household pre-harvest food consumption per capita	18712	61830.292	38998.913
Total household pre-harvest food consumption per capita (market)	18712	44371.909	34966.723
Total household pre-harvest food consumption per capita (self)	18712	15222.301	13071.153
Food insecure (pre-harvest)	18712	0.355	0.478

Notes: Descriptive statistics at the origin-destination pair year level. The sample period includes all years 2011 and 2014 to 2019. Seasonal migration variables are defined for localities for which the maximum value is greater than 0, which are the observations contributing to estimation with the PPML estimation. Departure in temporary or permanent migration are dummy variables equal to one if an individual leaves the houshold in the third or fourth quarter of data collection. Temporary or permanent nature of departure is reported by the household of origin. Conflict variables are defined in terms of count of fatalities and dummy if any fatality occured and measured between the first and second quarter of data collection. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance to closest city is measured in euclidian distance.

contributing observations used by the PPML models depending on the structure of fixed effects.[7] The last panel shows the descriptive statistics of the dyadic panel weighted by the number of migrants for each pair so that the figures are representative of the migration conditions faced by the average economic migrant.

5 Empirical strategy

5.1 Assessing the effect of route conflicts on bilateral economic migration and remittances

I estimate equation 4 using with the ratio of economic migrants to adult household members left-behind as empirical counterpart of $\frac{m_{odt}}{m_{oot}}$. Following the migration literature, I assume that bilateral migration costs c_{odt} are a linear combination of the log of distance between o and d Ω_{od} and the share of individuals at destination d whose mother tongue is the origin's main language, π_{od}. The attractiveness of origin o is a function of the log of origin wage ω_{ot} and destination attractiveness is similarly defined as a function of destination's wages; Origin fixed effects η_o and destination fixed effects θ_d included in all specifications control for the effect of time-invariant counfonding factors such as development level, population, institutions, culture, attitudes toward migration that can determine migration and conflict, while year fixed effects δ_t capture all time-varying terms that are constant across origins and destinations. The estimating equation also includes well-behaved error term ϵ_{odt}, with $E\left(\epsilon_{odt}\right) = 1$.

$$\frac{m_{odt}}{m_{oot}} = \exp\left\{\alpha_1 \log(\Omega_{od}) + \alpha_2 \log(\pi_{od}) + \alpha_3 \log(\omega_{ot}) + \alpha_4 \log(\omega_{dt}) + \delta_t + \eta_o + \theta_d\right\} \cdot \epsilon_{odt} \tag{5}$$

I extend this classical gravity model of migration adding measures of origin, destination, and route conflict, with μ_{ot-1} the count of conflict events at the origin o, μ_{dt-1} conflict at destination, and μroad_{odt-1} conflicts on the route between o and d.

$$\begin{aligned}\frac{m_{odt}}{m_{oot}} = \exp\{&\alpha_1 \log(\Omega_{od}) + \alpha_2 \log(\pi_{od}) + \alpha_3 \log(\omega_{ot}) + \alpha_4 \log(\omega_{dt}) \\ &+ \kappa_1 \log(\mu_{\text{road}_{odt-1}}) + \kappa_2 \log(\mu_{dt-1}) + \kappa_3 \log(\mu_{ot-1}) \\ &+ \delta_t + \eta_o + \theta_d\} \cdot \epsilon_{odt}\end{aligned} \tag{6}$$

Alternatively, in a second set of specifications I include origin-by-year fixed effect as in Ortega and Peri (2021b) to control for time-varying conditions of origin that could affect migration levels, including multilateral resistance to migration. This implies that overall reduction in migration propensity due to reduction in the average accessibility of destinations, including due to conflict, will be absorbed by the origin-by-year fixed effect. This specification further allows to relax the assumption of perfect substituability between origin o and all other locations $D = L \setminus \{o\}$, and is consistent with a nested logit model where migrants differ systematically from stayers (Ortega and Peri, 2013).

[7] Since the PPML model keeps only contributing observations in the sample, which varies with the structure of fixed effect. For example including destination-by-year and origin-by-year fixed effects results in excluding all observations corresponding to origins or destinations in years with no migration, whereas these observations contribute to the estimation of year and locality fixed effect in the more parsimonious specification.

Table 4: Descriptive statistics dyadic panel

	N	mean	sd
	Panel A: Localities-years with at least one migrant		
	Sample origin-by-year and dest-by-year FE		
Internal migration corridors			
Internal migrants (any)	15294	0.205	0.404
Internal migrants/stayers	15294	0.0182	0.0674
Distance (km)	15294	683.4	442.8
Orig group at dest (share)	15294	0.112	0.193
Dest wages	15294	34016.3	12702.7
Orig wages	15294	29598.7	18225.5
Orig SPEI drought frequency	15294	0.409	0.805
Dest SPEI drought frequency	15294	0.389	0.741
Dest conflicts	15294	2.110	4.441
Any route conflict	15294	0.424	0.494
Route conflicts	15294	0.972	1.664
Orig conflicts	15294	0.168	0.677
	Panel B: Localities-years with at least one migrant sending remittances		
	Sample origin-by-year and dest-by-year FE		
Internal migration corridors			
Remittances (any)	14770	0.0371	0.0988
Remittances (amount)	4737	4879.7	9215.2
Distance (km)	14770	680.9	447.8
Orig group at dest (share)	14770	0.111	0.196
Dest wages	14770	33924.3	12873.0
Orig wages	13651	30292.4	17852.6
Orig SPEI drought frequency	14770	0.362	0.753
Dest SPEI drought frequency	14770	0.336	0.690
Dest conflicts	14770	2.856	5.163
Any route conflict	14770	0.452	0.498
Route conflicts	14770	1.629	2.979
Orig conflicts	14770	0.207	0.836
All migration corridors			
Remittances (any)	25671	0.0289	0.0882
Remittances (amount)	7268	4127.7	8471.3
Distance (km)	25671	820.6	541.8
Orig SPEI drought frequency	25671	0.375	0.778
Dest SPEI drought frequency	25671	0.444	0.862
Dest conflicts	25671	2.908	5.169
Any route conflict	25671	0.448	0.497
Route conflicts	25671	1.442	2.662
Orig conflicts	25671	0.191	0.797
	Panel C: Pair-years with at least one migrant		
	Weighted by number of migrants		
Internal migration corridors			
Internal migrants (any)	3143	1.000	0.00299
Internal migrants/stayers	3143	0.278	0.402
Remittances (any)	3143	0.174	0.195
Remittances (amount)	2877	7493.6	12789.0
Distance (km)	3143	476.1	393.0
Orig group at dest (share)	3143	0.113	0.205
Dest wages	3143	34485.2	12050.4
Orig wages	3143	26782.4	19053.7
Orig SPEI drought frequency	3143	0.370	0.779
Dest SPEI drought frequency	3143	0.432	0.796
Dest conflicts	3143	0.574	1.906
Any route conflict	3143	0.237	0.426
Route conflicts	3143	0.300	0.724
Orig conflicts	3143	0.0623	0.340

Notes: Descriptive statistics at the orign-destination pair level. The sample period includes all years 2011 and 2013 to 2019. The different panel correspond to samples with observations contributing to the estimation with the PPML estimation depending on the fixed effects included in the regression. Internal migrants is the average number of migrants reported by the households in the locality for the regions of Mali, while remittances (any) is the share of household reporting receiving remittances from the set of twelve destinations, and remittances (amount) the average amount received by households in the locality. Conflict variables are defined in terms of count of fatalities and measured six month prior to the average regional date of harvest in the locality of origin. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest road between the two.

Additionally, destination-by-year fixed effects address time-varying destination-specific confounding factors, such as economic volatility at destination. This is also important given that measurement error in fatalities at destination, due to the lower geographical precision of the actual place of destination (regional or country level), could result in the bilateral conflict cost terms to capture conditions at destinations. With this structure of fixed effects, only the effect of pair-specific characteristics X_{od} can be estimated. While the sample size and relatively high number of pairs with no migration flows does not allow to control simultaneously for pair fixed effect destination-by-year and origin-by-year fixed effects, I add to the last specification the share of preconflict migrants between to the set of pair specific characteritics X_{od} to control for potential confonding factors specific to a given migration corridors, such as the quality of transportation infrastructure.

$$\frac{m_{odt}}{m_{oot}} = \exp\left\{\boldsymbol{\alpha X_{odt}} + \kappa_1 \log(\mu_{\text{road}_{odt-1}}) + \delta_t + \eta_o + \theta_d + \gamma_{dt} + \phi_{ot}\right\} \cdot \epsilon_{odt} \tag{7}$$

Equations 5 to 7 are estimated directly with Poisson Pseudo Maximum-Likelihood (PPML), with robust standard errors clustered by origin-destination pair. While the theoretically driven estimation equation refers to migration flows, empirical applications of the gravity models have alternatively used measures flows and stocks of migrants (Mayda, 2010). My main variable of interest m_{odt} is the average stock of migrants per household from origin o, currently residing in d. To shed light on potential welfare implications of reduced migration flows, alternative specifications replace the average number of migrants by the share of households receiving remittances and the amount of remittances received. The relationship between remittances and migration is not mechanical and independent from conflict, as conflict fatalities could affect the decision of migrants to transfer money back to their household of origin, who could increase their remittances to compensate for reduced number of migrants, or to substitute for direct transfers upon their return. Unless these compensating mechanisms fully offset the effect of reduced migration, I expect the sign of the coefficient of the effect of route conflicts on the share of household receiving remittances, and the amount received, to be negative.

With the rich structure of fixed effects included in the empirical specification, the main concern to identification remains reverse causality: more migration on a specific corridor could lead to higher number of fatalities if migrants are victims of conflicts. This potential mechanism, resulting in a positive realtionship between migration and conflict, would bias downwards the estimates. To mitigate this concern, in supplementary analysis, the number of fatalities is instrumented by temporally and spatially lagged values of conflict. This instrument is constructed in a stage zero grid-level regression, with 10646 cells of 0.1° (approximatively 10km2), where the number of conflict fatalities in a given cell is predicted by conflicts in the neighboring cell in the past year. The final value of the instrument is the average predicted number of fatalities in cells crossed by the shortest road between each locality and destination, weighted by the cell area within the 2km buffer area around the road.

5.2 Assessing the effect of migration routes conflicts on households members mobility and household welfare.

I construct a measure of average route conflicts relevant to each locality *migration market* as the average number of conflicts on the route between each locality and destination weighted by its pre-conflict share of migrants.

$$\overline{\mu\text{road}_{ot}} = \left[\sum_{d\in D} \mu\text{road}_{odt} \cdot \frac{m_{od_{2009}}}{m_{oo_{2009}}}\right] \cdot \left[\sum_{d\in D} \frac{m_{od_{2009}}}{m_{oo_{2009}}}\right]^{-1} \tag{8}$$

I identify the pre-conflict share of migrants $\frac{m_{od_{2009}}}{m_{oo_{2009}}}$ for each locality o and destination d pair using the 2009 population census that measures past destinations of return migrants. The gravity analysis reveals that this share is strongly correlated with economic migration during the period of interest. Figure A 1 in appendix represents the average exposure to conflict on migration routes for each surveyed locality by year of data collection. The same variable is contructed for the average conflicts at destination.

The average route and destination conflict variable are built as exogenous shocks to migration, similar to the shift-share instruments used in the migration and trade literature. These instruments utilize shocks to destinations that affect the unit of interest differently depending on their share of migrants or exports to these destinations. However, the average migration route insecurity adds an additional layer of variation depending on the spatial location of localities and each destination. To ensure that destination shocks are exogenous to the locality conditions, I exclude the own region from the set of possible destinations.

I estimate the effect of this route and destination insecurity on three sets of variables with different objectives in mind.

Firstly, by aggregating route conflicts across all potential migration destinations, I am able to estimate the cumulative impact of route insecurity on migration. While the gravity equation allows for a more precise identification of the effect of insecurity on a specific migration corridor, reduced migration to one destination might be compensated for by a shift towards other destinations. This redirection of migration flows to alternative destinations carries significant welfare implications, as it suggests that migrants may not be able to select their optimal destinations in terms of expected returns and migration costs in the absence of conflict. It is however important to assess whether this phenomenon is associated with an overall reduction in migration flows.

The following econometric model, estimated with PPML, reproduces the baseline gravity estimation aggregated at the locality level:

$$\frac{m_{ot}}{m_{oot}} = \exp\left\{\alpha_0 + \kappa_1 \log(\overline{\mu\text{road}_{ot}}) + \kappa_2 \log(\overline{\mu_{dt}}) + \kappa_3 \log(\mu_{ot}) + \delta_t + \eta_c\right\} \cdot \epsilon_{ot} \tag{9}$$

where m_{ot} refers alternatively to the three migration variables defined in section 5.1: the ratio of migrants to all internal destination to adult household members left-behind, the average share of households receiving remittances across the set of potential destinations, and the total amount of remittances received from the set of potential destinations; μ_{ot-1} is a count variable of the number of conflict events occuring in a given 50 km buffer around locality c in the pre-harvest period, $\overline{\mu\text{road}_{ot-1}}$ the average number of route conflicts in the same period and $\overline{\mu_{dt-1}}$ the average destination conflict intensity calculated with the same weights; δ_t and η_c are year and locality fixed effects. To account for the correlation of observations within localities, I cluster standard errors at the locality and year level.

The second goal of an aggregated estimation is to use an alternative migration variable that cannot be disagregated over destinations with the data currently at hand but provides an objective and precisely dated measure of migration flows: household members departures between quarters of survey. This estimation allows me to leverage the calendar of data collection that spans over the agricultural cycle: most migration departures occur in the post-harvest season, and households are surveyed twice in the pre-harvest season, and then again twice in the post-harvest season.

Hence using the date of interview for each locality and year, I can isolate the effect of conflict events prior to potential post-harvest migration departures (i.e. the date of the second pre-harvest data collection survey), on post-harvest departures. This high frequency flow data therefore allows to mitigate potential reverse causality mechanisms, where migration would be the reason for an increase number of victims of conflicts on migration roads, or alternative mechanisms including reduced return migrants inflows or increased migrants mortality, rather than reduction of migrations outflows. Furthermore, this alternative measure of migration flows includes not only economic migration but also other forms of migration strategies. Existing literature has demonstrated that risk-coping migration extends beyond labor-related migration and encompasses practices such as child fostering and the marriage of young women. To test for the response of migration flows to average road insecurity, I estimate equation 9 with m_{iot}, the propensity to engage in post-harvest migration for individual i in locality o at time t as a dependent variable.

Finally, this specification is used to estimate the aggregate impact of route insecurity on household consumption. I focus on household consumption per capita C_{hot} during the pre-harvest season as an indicator of households well-being.

6 Results

6.1 Route conflicts and bilateral economic migration

Table 5 displays the results of the gravity model for internal migration. The explanatory variables in the first column are the inverse hyperbolic sine transformation of the average wage at destination and origin, the ihs transformation of distance, and the ihs transformation of the share of individuals at destination that report as mothertongue the origin main language. These variables have been shown to be important determinants of migration flows in the literature and display the expected sign, with migration increasing with destination average wage and shared language, and decrasing in wage at origin and distance, although only the coefficients for distance and destination wages are significant. The coefficients can be interpreted as elasticities such that a one percent change in the independent variable corresponds to an equivalent percentage change of the dependent variable equal to the size of the coefficient. For example, a one percent increase in distance to destination is associated with a 0.51 to 0.56 percent decrease in the number of internal migrants [8].

Columns 2 to 6 reveal significant and large elasticities of internal migration with respect to bilateral route conflict events, with estimated coefficients between 0.44 and 0.89. The estimated coefficient in column 2 decreases when including controls for traditional determinants of migration in column 3, but increases when including origin-by-year and destination-by-year fixed effects in columns 5 and 6. This means that once overall effect of time-varying shocks affecting average migration levels from a given locality are absorbed, destination choices of migrants are constrained by route conflicts. In the migration (and trade) literature, these fixed effects have been shown to capture the *multilateral resistance* to migration (or trade), and absorb variations in average migration levels. As such, while the coefficients for route conflicts in columns 3 and 4 account

[8] This elasticity to distance is lower than the one traditionally found in the literature on international migration that is generally around 0.7. This may be due to the fact that the analysis focuses on regional migration, and at this level, equally distant destinations can uncover highly heterogenous transportation costs depending on the quality of infrastructure, even when distance refers to the actual road distance on the shortest route.

Table 5: Gravity model of internal migration

	Internal migrants/Stayers					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance (ihs)	-0.556***		-0.513***	-0.514***		
	(0.169)		(0.143)	(0.142)		
Share orig. group at dest. (ihs)	0.255		0.207	0.204		
	(0.235)		(0.224)	(0.226)		
Dest. av. wage (ihs)	0.867**		0.477	0.446		
	(0.429)		(0.357)	(0.378)		
Orig. av. wage (ihs)	-0.049		-0.051	-0.049		
	(0.033)		(0.057)	(0.060)		
Route conflicts (ihs), t-1		-0.624*	-0.437**	-0.433**	-0.890***	-0.723***
		(0.330)	(0.208)	(0.207)	(0.239)	(0.206)
Orig. conflicts (ihs), t-1		0.157**	0.147*	0.145*		
		(0.061)	(0.089)	(0.087)		
Dest. conflicts (ihs), t-1		-0.159***	-0.164**	-0.167**		
		(0.061)	(0.065)	(0.071)		
Orig. SPEI drought frequency (ihs)				-0.026		
				(0.064)		
Dest. SPEI drought frequency (ihs)				0.079		
				(0.157)		
Share pre-conflict return migrants (ihs)						2.39***
						(0.261)
P-values route coef.						
SE clustered by year and locality		0.059	0.036	0.037	0.0002	0.0004
Conley SE		2.11×10^{-8}	1.6×10^{-6}	2.08×10^{-6}	1.48×10^{-11}	2.58×10^{-9}
Pseudo R^2	0.2748	0.2658	0.2779	0.2779	0.3057	0.3154
Observations	15,294	15,294	15,294	15,294	15,294	15,294
Dependent variable mean	0.0182	0.0182	0.0182	0.0182	0.0182	0.0182
Implied effect for one additional conflict		-0.012	-0.008	-0.008	-0.017	-0.014
Year fixed effects	✓	✓	✓	✓		
Destination fixed effects	✓	✓	✓	✓		
Locality fixed effects	✓	✓	✓	✓		
Locality-Year fixed effects					✓	✓
Destination-Year fixed effects					✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The p-values displayed at the bottom of the table are only for the route conflict coefficients, and conley standard errors allow fo spatial correlations up to 50km. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

for a decrease in migration levels due to conflicts-induced reduction in accessibility of potential destinations, this effect is absorbed by the locality-by-year fixed effect in columns 5 and 6.

The larger coefficients in specifications controlling for origin-by-year and destination-by-year fixed effects in columns 5 and 6 may be due to omitted variables correlated with route conflicts and economic migration, that would go in the opposite direction as the observable characteristics included as controls in the second column. However, these larger coeffcents can also suggest that substitution between migration destinations partially offsets the reduction of migration on corridors affected by insecurity.

While the negative effect of route insecurity on economic migration is constantly significant across specifications, its magnitude varies largely depending on whether the fixed effects absorb the overall conflict-induced reduction in migration. As such, the highest estimated elasticity of 0.89 in column 5 accounting only for substitution between migration destinations implies that one additional route conflict is associated with a reduction of 0.017 of the average migrant to stayer ratio for a given migration corridor, meaning 0.153 migrant for an average household of 9 members. Instead, the estimated elasticity in the absence of year specific locality and destination fixed effects would imply a reduction of 0.008 internal migrants per household member.

Table A6 in appendix estimates the same regressions including origin-by-destination pair fixed effects on a subsample of migration corridors with at least one migrant in one year. While this restriction greatly reduces the sample size and sample restrictions do not allow to control simultaneously for pair fixed effects, destination-by-year and locality-by-year fixed effects simultaneously, the estimated elasticity is stable with the inclusion of the pair fixed-effect, suggesting that ommitted characteristics are not driving the effect.[9]

While causality is not established for these coefficients, specifications that do not incorporate origin-by-year fixed effects allow us to examine the impact of conflict at the origin on migration. This effect is positive and significant in all estimated specifications, suggesting that insecurity may act as a push factor for economic migrants. Conversely, the impact of destination conflicts can only be estimated in specifications where the year-to-year variation in destination characteristics is not absorbed by fixed effects. Notably, this effect is consistently negative and statistically significant across all specifications where it is included.

Table 6 displays the results of the gravity model for the receipt of internal remittances on the intensive and extensive margins, focusing on the effect of conflict variables. It is important to note that while the theoretical framework does not explicitly model the decision to remit conditional on migration decisions, this choice may not be orthogonal to migration costs. While most remittances are now tranferred in Mali through mobile banking (EMOP 2022), this was probably less the case at the start of the period. Additionally, remittances may serve as a substitute for other more direct forms of contributions from migrants to household expenses, particularly those that involve traveling back to the locality of origin. The elasticity of the share of households receiving remittances from a given destination as well as the amount is negative and signfication across all specifications. The estimated results that for one additional conflict one a given migration corridor, households are on average between 0.8 and one percent less likely to receive remittances from this given destination, which correspond to an average loss of around 3000 FCFA over a year, about a fifth of average monthly wages in the sample. The substantial positive impact of drought in the locality of origin on both the share of households receiving remittances and the average amount of transfers received underscores the role of migration as an insurance mechanism. Together with the

[9]Note also that the estimated effect is an elasticity and that the mean of the dependent variable is larger in this subsample while the mean of route conflicts is lower, so that the unit effect is of similar magnitude.

null effect of drought on the migrant to stayers ratio estimated in Table 5, these results suggest that a drought at the origin leads to an increased propensity among migrants to remit their income back home rather than causing an increase in the number of economic migrants, which is consistent with liquidity constraints preventing costly migration.

Table 6: Gravity model of internal remittances

	Any transfer				Amount received (FCFA)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Route conflicts (ihs), t-1	-0.351***	-0.342***	-0.440***	-0.366***	-0.462***	-0.462***	-0.557***	-0.475***
	(0.116)	(0.117)	(0.088)	(0.075)	(0.126)	(0.134)	(0.144)	(0.127)
Orig. conflicts (ihs), t-1	0.076***	0.058			0.071***	0.063**		
	(0.029)	(0.051)			(0.015)	(0.030)		
Dest. conflicts (ihs), t-1	-0.071	-0.100			0.169	0.049		
	(0.101)	(0.113)			(0.283)	(0.354)		
Orig. SPEI drought frequency (ihs)		0.181***				0.116***		
		(0.023)				(0.037)		
Dest. SPEI drought frequency (ihs)		0.197				0.158		
		(0.138)				(0.249)		
Share pre-conflict return migrants (ihs)				1.41***				1.17***
				(0.238)				(0.349)
Pseudo R^2	0.2893	0.2904	0.3273	0.3308	1.082	1.082	1.060	1.059
Observations	13,651	13,651	13,651	13,651	3,624	3,624	3,624	3,624
Dependent variable mean	0.0373	0.0373	0.0373	0.0373	4,787.7	4,787.7	4,787.7	4,787.7
Implied effect for one additional conflict	-0.008	-0.008	-0.010	-0.008	-2,874.3	-2,873.1	-3,466.2	-2,958.3
Year fixed effects	✓	✓			✓	✓		
Destination fixed effects	✓	✓			✓	✓		
Locality fixed effects	✓	✓			✓	✓		
Locality-Year fixed effects			✓	✓			✓	✓
Destination-Year fixed effects			✓	✓			✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The p-values displayed at the bottom of the table are only for the route conflict coefficients, and conley standard errors allow fo spatial correlations up to 50km. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

The results displayed in Table 7 suggest that road conflicts constrain migration decisions even more in localities hit by a negative shock: while an increase in the frequency of droughts during the past growing season is associated with higher reported stock of internal migrants, a higer share of households receiving remittances and larger amounts received, the migratory response to such events is lower on migration corridors affected by conflicts. These results align with the hypothesis of heightened liquidity constraints in areas impacted by droughts, leading to increased sensitivity of households to migration costs.

After establishing that conflicts along migration routes do reduce migration and remittances on affected migration corridors, I turn to explore the aggregate effect of increased route insecurity and reduced migration access on economic migration and household members mobility as well as its welfare implications.

Table 7: Gravity model of internal migration and remittances by drought at origin

	Internal migrants/Stayers		Any transfer		Amount received (FCFA)	
	(1)	(2)	(3)	(4)	(5)	(6)
Route conflicts (ihs), t-1	-0.890***	-0.670***	-0.440***	-0.431***	-0.557***	-0.432***
	(0.239)	(0.250)	(0.088)	(0.089)	(0.144)	(0.115)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency (ihs)		-0.431***		-0.039		-0.331*
		(0.159)		(0.161)		(0.190)
Pseudo R^2	0.3057	0.3067	0.3273	0.3273	1.060	1.059
Observations	15,294	15,294	13,651	13,651	3,624	3,624
Dependent variable mean	0.0182	0.0182	0.0373	0.0373	4,787.7	4,787.7
Locality-Year fixed effects	✓	✓	✓	✓	✓	✓
Destination-Year fixed effects	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

6.2 Aggregate effects of route conflict on migration

The first six columns of table 8 report the results of the aggregated analysis using a PPML estimation for the same variables used in the gravity model, including the ratio of economic migrants to adult household members left-behind, the average share of households receiving remittances, and average amount of remittances. Although no significant effect of route conflict can be detected for these variables, for internal migration the standard errors do not exclude that effects as large as those estimated with the gravity model cannot be precisely estimated. As such, the results for economic migration are inconclusive to whether any reduction in migration along specific routes due to insecurity is fully compensated by increased migration in other corridors. Similar to dyadic regressions, the ratio of migrants to stayers does not exhibit a significant increase with drought frequency, although the coefficient is positive. In contrast, both the share of households receiving remittances and the amount received register substantial increases in response to drought shocks.

The last four columns of table 8 focus on potential implications of substitution between migration destinations, namely whether migrants faced with insecurity on their usual migration destinations turn to closer destinations, or destinations with lower average wage. Although the sign of the route conflict coefficient is negative for both the average distance to destination and average wage at destination, it is only marginally significant for the average distance to destination. The estimated effect implies that economic migrants facing one additional conflict on their usual migration routes choose destinations on average 58.6 km closer to their locality of origin. Interestingly, the same effect can be observed for migrants from localities affected by droughts, suggesting that similar liquidity constraints mecanisms restrict the choice set of economic migrants in case of drought and route insecurity.

Table 9 presents the effect of average route conflicts on household members mobility, including post-harvest departures and early season returns.

The elasticity of permanent and temporary departures with respect to route conflicts is not significant on average (columns 1, 3 and 5). Temporary departures of household members however, are significanly reduced by route conflicts for individuals in localities affected by droughts in the pre-harvest growing season. Post-harvest departures increase on average with the frequency of drought during the growing season, and this effect is mostly driven by permanent departures suggesting that permanent departure of a household member is a risk-coping strategy for households.

Table 8: Average route insecurity and economic migration

	Economic migrants in Mali									
	Internal migrants/Stayers		Any transfer		Amount (FCFA)		Av. dest. wage		Av. dest. dist.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs), t-1	-0.246	-0.362**	-0.066	-0.070	-0.053	-0.082	-0.024	-0.012	-0.083*	-0.075
	(0.151)	(0.180)	(0.114)	(0.120)	(0.097)	(0.108)	(0.029)	(0.030)	(0.044)	(0.050)
Orig. conflict (ihs), t-1	0.084	0.109	-0.009	-0.008	-0.009	-0.0009	-0.041	-0.043*	0.052	0.050
	(0.131)	(0.128)	(0.100)	(0.101)	(0.096)	(0.096)	(0.026)	(0.025)	(0.053)	(0.053)
Orig. SPEI drought frequency, t-1	0.034	-0.114	0.322***	0.314***	0.296***	0.252***	-0.011	0.0002	-0.075***	-0.065**
	(0.103)	(0.124)	(0.084)	(0.108)	(0.071)	(0.091)	(0.012)	(0.014)	(0.029)	(0.030)
Dest. conflict (ihs), t-1	0.132	0.153	-0.071	-0.071	-0.114	-0.108	-0.038**	-0.040**	0.018	0.017
	(0.115)	(0.111)	(0.062)	(0.062)	(0.078)	(0.079)	(0.017)	(0.017)	(0.043)	(0.044)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency, t-1		0.278**		0.018		0.097		-0.023		-0.018
		(0.114)		(0.077)		(0.080)		(0.014)		(0.032)
Pseudo R^2	0.1468	0.1475	0.1488	0.1488	1.211	1.211	1.001	1.001	1.013	1.013
Observations	14,901	14,901	16,956	16,956	16,323	16,323	10,675	10,675	10,675	10,675
Dependent variable mean	0.0847	0.0847	0.0155	0.0155	8,046.3	8,046.3	37,231.1	37,231.1	452.2	452.2
Implied effect for one additional conflict (SPEI=0)	-0.032	-0.048	-0.002	-0.002	-689.3	-1,064.0	-1,402.0	-691.6	-58.6	-52.9
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

Conversly, conflicts at destination reduce the probability of permanent departures only. Attrition of households in the post-harvest quarters does not seem to be correlated with road conflicts, but increases instead significantly with conflicts at origin. The probability of a member to return is not significantly correlated with any for the independent variables, although this mesure is imperfect as both likelihood of leaving and returning to the household will be affected by insecurity.

Table 9: Average route insecurity and households mobility

	Post-harvest departures								Early season returns	
	All departures		Temporary		Permanent		Attrition		Return mig. in q1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs)	-0.221	-0.153	-0.211	-0.084	-0.209	-0.179	0.166	0.222	-0.121	-0.092
	(0.140)	(0.152)	(0.167)	(0.141)	(0.186)	(0.246)	(0.359)	(0.436)	(0.138)	(0.202)
Orig. conflict (ihs)	0.166	0.156	0.127	0.108	0.290	0.286	2.17***	2.17***	0.038	0.035
	(0.118)	(0.120)	(0.109)	(0.109)	(0.255)	(0.256)	(0.716)	(0.726)	(0.176)	(0.173)
Dest. conflict (ihs)	-0.080	-0.084	0.033	0.021	-0.394***	-0.394***	-0.180	-0.186	-0.053	-0.056
	(0.053)	(0.052)	(0.085)	(0.082)	(0.120)	(0.120)	(0.369)	(0.370)	(0.087)	(0.088)
Orig SPEI drought frequency	0.162*	0.186**	0.084	0.126	0.391***	0.403***	0.024	0.061	0.087	0.095
	(0.083)	(0.084)	(0.089)	(0.102)	(0.086)	(0.075)	(0.227)	(0.197)	(0.086)	(0.089)
Route conflicts (ihs) × Orig SPEI drought frequency		-0.140		-0.298***		-0.052		-0.131		-0.048
		(0.099)		(0.107)		(0.193)		(0.444)		(0.140)
Pseudo R^2	0.0666	0.0667	0.0740	0.0744	0.0911	0.0912	0.2431	0.2432	0.1053	0.1053
Observations	180,709	180,709	176,351	176,351	134,758	134,758	4,776	4,776	169,460	169,460
Dependent variable mean	0.0518	0.0518	0.0388	0.0388	0.0186	0.0186	0.0773	0.0773	0.0536	0.0536
Implied effect for one additional conflict (SPEI=0)	-0.0229	-0.0158	-0.0167	-0.0066	-0.0077	-0.0065				
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

These results are not necessarily inconsistent with the economic migration results in table 8 as the migration captured in the economic migration module is essentially economically motivated, while post-harvest departures could emcompass any type of mobility. To investigate further the response of these different types of mobility, I further study the effect of route conflicts on post-harvest depatures by gender in Table 10, further detailed by gender and age groups in Table A2 in appendix. While there is an average significant reduction of post-harvest departures for women, for men this effect is significant only when interacted with droughts. Further looking into age groups, it appears that the negative average effect of route is significant for both gender when considering children below 14 years old. The results by age group are also insightful to understand households' response to droughts : while for adult members drought at origin increases post-harvest departures of men and women, the same shocks affects boys only, suggesting gender-specific child fostering (which is opposite to Akresh (2009) results in Burkina Faso showing increased fostering of girls for households facing income shocks).

Table 10: Average route insecurity and post-harvest departures by gender

	Men						**Women**					
	All departures		Temporary		Permanent		All departures		Temporary		Permanent	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Route conflicts (ihs)	-0.180	-0.104	-0.160	-0.025	-0.218	-0.138	-0.282*	-0.229	-0.290	-0.178	-0.186	-0.249
	(0.153)	(0.165)	(0.196)	(0.159)	(0.222)	(0.269)	(0.149)	(0.176)	(0.182)	(0.197)	(0.190)	(0.260)
Orig. conflict (ihs)	0.186	0.175	0.174	0.152	0.252	0.244	0.130	0.123	0.051	0.036	0.347	0.360
	(0.130)	(0.130)	(0.130)	(0.126)	(0.288)	(0.286)	(0.119)	(0.123)	(0.120)	(0.122)	(0.247)	(0.254)
Dest. conflict (ihs)	-0.134**	-0.138**	-0.058	-0.071	-0.328**	-0.325**	-0.009	-0.013	0.142	0.130	-0.477**	-0.476**
	(0.059)	(0.057)	(0.084)	(0.083)	(0.133)	(0.128)	(0.076)	(0.076)	(0.131)	(0.129)	(0.214)	(0.211)
Orig SPEI drought frequency	0.159*	0.187**	0.052	0.098	0.441***	0.476***	0.174**	0.191**	0.128	0.165	0.327**	0.304***
	(0.092)	(0.092)	(0.108)	(0.120)	(0.086)	(0.084)	(0.086)	(0.092)	(0.093)	(0.107)	(0.137)	(0.117)
Route conflicts (ihs) × Orig SPEI drought frequency		-0.160		-0.320***		-0.142		-0.103		-0.257*		0.099
		(0.114)		(0.106)		(0.229)		(0.112)		(0.147)		(0.188)
Pseudo R^2	0.0698	0.0700	0.0706	0.0711	0.0899	0.0901	0.0696	0.0697	0.0726	0.0729	0.0903	0.0903
Observations	86,463	86,463	79,572	79,572	63,026	63,026	86,154	86,154	82,956	82,956	49,751	49,751
Dependent variable mean	0.0604	0.0604	0.0465	0.0465	0.0242	0.0242	0.0479	0.0479	0.0379	0.0379	0.0197	0.0197
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

To further shed light on the average effect of route insecurity and its interaction with droughts, I split the sample between households in the bottom and top half of the distribution a wealthindex created based on housing characteristics, in Table 11. For the poorest household of the sample, there is no average effect of route insecurity on post-harvest departures, but a significant effect of the interaction of route conflicts and droughts. Instead, for the richest household, route insecurity signifcantly reduces post-harvest depatures on average, but there is no additional effect of the interaction with droughts. These differential responses to route conflicts may suggest that avoidance of conflict may dominate for women, children and richest households, while a liquidity constraint channel may explain the reduction of post-harvest departures for adult men and poorest households faced both with route insecurity and droughts.

Table 11: Average route insecurity and post-harvest departures by wealth

	Poorest Households						**Richest Households**					
	All departures		Temporary		Permanent		All departures		Temporary		Permanent	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Route conflicts (ihs)	-0.178	-0.033	-0.085	0.109	-0.326	-0.216	-0.304**	-0.323*	-0.371**	-0.284	-0.161	-0.295
	(0.196)	(0.177)	(0.247)	(0.206)	(0.202)	(0.263)	(0.127)	(0.175)	(0.162)	(0.189)	(0.203)	(0.258)
Orig. conflict (ihs)	0.264*	0.245*	0.292**	0.274**	0.267	0.239	0.101	0.105	-0.030	-0.047	0.448	0.477
	(0.141)	(0.142)	(0.139)	(0.137)	(0.296)	(0.297)	(0.150)	(0.155)	(0.143)	(0.149)	(0.314)	(0.322)
Dest. conflict (ihs)	-0.038	-0.053	0.111	0.086	-0.397***	-0.405***	-0.072	-0.072	-0.029	-0.035	-0.203	-0.207
	(0.078)	(0.071)	(0.131)	(0.125)	(0.135)	(0.139)	(0.124)	(0.124)	(0.164)	(0.163)	(0.189)	(0.190)
Orig SPEI drought frequency	0.111	0.176	0.100	0.179	0.260	0.321*	0.170*	0.164	0.088	0.116	0.413***	0.372***
	(0.127)	(0.134)	(0.118)	(0.132)	(0.180)	(0.184)	(0.101)	(0.103)	(0.117)	(0.131)	(0.097)	(0.092)
Route conflicts (ihs) × Orig SPEI drought frequency		-0.330**		-0.467***		-0.225		0.034		-0.197		0.178
		(0.137)		(0.145)		(0.336)		(0.110)		(0.128)		(0.182)
Pseudo R^2	0.0782	0.0788	0.0829	0.0839	0.0899	0.0902	0.0732	0.0732	0.0793	0.0795	0.1030	0.1032
Observations	82,429	82,429	73,982	73,982	51,760	51,760	87,601	87,601	84,295	84,295	56,898	56,898
Dependent variable mean	0.0529	0.0529	0.0425	0.0425	0.0236	0.0236	0.0570	0.0570	0.0439	0.0439	0.0227	0.0227
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

6.3 Aggregate effects of road insecurity on household consumption

The results presented in section 6.1 and 6.2 point towards a substitution of economic migration to safely accessible destinations and a reduction in aggregate seasonal migration departures associated with road insecurity in localities affected by episodes of droughts. Although I find no evidence of aggregate reduction in remittances associated with increased road insecurity, this does not exclude that the reduction and reallocation of migration flows carries welfare implications. Firstly, the aggregate results are less well-identified than the gravity model results, so that we cannot exclude a reduction of transfers due to increased road insecurity. Second, even in the absence of reported transfers, direct contributions of migrants can take many forms that may not be labelled and reported as transfers, such as in-kind transfers, direct transmission of assets upon their return, or bride price in the case of permanent departure for marriage. Finally, migration of a household member in case of severe shock may contribute to sustain a minimum household consumption per capita through adjutsment in household size and composition, for example through child fostering as documented in the analysis of household members mobility.

I investigate the net effect of route insecurity on household welfare by focusing on consumption expenditures during the lean season, the season during which rural incomes and harvest stocks are at their lowest point, and hence income from external sources such as migration will matter most. Since consumption expenditures are measured in local currency and data on prices are not available for the 2011-2019 survey period, I do not aim to estimate the exact magnitude of the effect of conflict on real consumption expenditures but rather provide some lower bounds of this effect in the absence of price elasticity. While I cannot completely disentangle migration networks from trade networks, all specifications control for the current values of the variables of interest which could affect households ability to purchase goods, while I will be focusing my analysis on the impact of route conflicts during the past year pre-harvest period, which, has shown in section 6.2, would have affected household capacity to adjust its composition during the past post-harvest season, and may hence affect current household consumption through this reduced migration channel. The results presented in table 12 suggest that conflicts on migration routes are associated with lower nominal per capita consumption expenditures during the pre-harvest quarters, with one additional route conflict corresponding to a reduction of about 8720 FCFA, about 14% of consumption for the period. On the other hand, in localities hit by droughts during the growing season, route conflicts are associated with higher food insecurity. The absence of average effect of route conflict on food insecurity may be explain by a partial compensation of households of reduced expenditures for food purchased on markets through increased self-consumption, which is not available in case of droughts during the past growing season, as shown in column 8.

6.4 Discussion

Identification of conflict as a bilateral migration cost. By focusing on specific migration corridors, the dyadic regressions allow to interact origin and destination fixed effects with year fixed effect, and hence to control for any time-varying shock affecting a given locality or potential migration destinations. This approach strengthens the argument for a causal interpretation of the estimates, as it reduces the potential bias from omitted variables that may impact both migration decisions and conflict outcomes. The main identification concern with this strategy remain reverse causality where migration flows between two locations would directly induce higher conflict fatalities on the route between these two locations. The results from the instrumental variable strategy in Table A5 , where road conflicts are instrumented by temporal and spatial lags of conflicts (i.e.,

Table 12: Average route insecurity : welfare implications

	Pre-harvest per capita expenditures (ihs)									
	Total		Food		Food (market)		Food (self)		Food insecure (q1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs), t-1	-0.066**	-0.062**	-0.060***	-0.054***	-0.085***	-0.087***	0.038	0.066**	0.044	-0.015
	(0.026)	(0.025)	(0.021)	(0.017)	(0.028)	(0.024)	(0.032)	(0.033)	(0.048)	(0.051)
Orig. conflict (ihs), t-1	0.032	0.033	0.034	0.035	0.064***	0.070***	-0.057	-0.059	0.052	0.059*
	(0.022)	(0.026)	(0.021)	(0.024)	(0.023)	(0.021)	(0.070)	(0.069)	(0.039)	(0.035)
Dest. conflict (ihs), t-1	0.044	0.023	0.029	0.013	-0.020	-0.031	0.141***	0.090***	0.017	0.098
	(0.035)	(0.024)	(0.034)	(0.023)	(0.035)	(0.027)	(0.047)	(0.035)	(0.054)	(0.061)
Orig. SPEI drought frequency, t-1	0.010	0.018	-0.006	0.014	0.005	0.018	-0.028	-0.002	0.059*	-0.020
	(0.019)	(0.017)	(0.019)	(0.020)	(0.017)	(0.020)	(0.050)	(0.055)	(0.032)	(0.049)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency, t-1		-0.018		-0.036		-0.019		-0.073**		0.144**
		(0.022)		(0.022)		(0.024)		(0.036)		(0.060)
Pseudo R^2	-278.2	-1,635.5	-278.9	-1,662.7	-255.1	-1,552.3	-310.1	-2,116.4	0.1188	0.1328
Observations	18,716	18,716	18,716	18,716	18,716	18,716	18,716	18,716	17,918	17,918
Dependent variable mean	87,023.7	87,023.7	61,740.3	61,740.3	43,942.2	43,942.2	15,882.5	15,882.5	0.3589	0.3589
Implied effect for one additional conflict (SPEI=0)	-8,719.7	-8,237.5	-5,632.4	-5,068.5	-5,698.8	-5,829.3	934.0	1,588.2	0.023	-0.008
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls for variable in t	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

conflict in neighboring cells in the previous years) alleviates this concern. The complementary analysis in section 6.2 using an alternative measure of migration that can be precisely dated further weakens potential reverse causality concerns.

Identification of effect of average route conflicts, average destination conflicts and conflicts at origin. In the simple panel analysis, the two-way fixed effects adress any time-invariant counfounding factors at the locality level. Yet, the literature on conflict determinants (Miguel et al., 2004; Dube and Vargas, 2013; Harari and Ferrara, 2018) suggests that income and conflict are jointly determined, while income is itself a major determinant of migration (Clemens, 2020; Bryan et al., 2014). As such, localities on similar trends could be affected by a shock determining conflict and migration level in the short-run.

I first address this concern by studying the effect of droughts on conflict as it is a very important potential counfounding factor in a agriculture-based economy and can also inform us on the potential impact of similar unobserved shocks. I reproduce Harari and Ferrara (2018) grid-level panel analysis of determinants of conflict with 0.5° squares and find no short-term effect of own-cell drought on conflict for Mali when controlling for drought in neighboring cells (see table A3 columns (5) to (8) in appendix), which is consistent with McGuirk and Nunn (2021) findings that conflict in Africa are determined by droughts in neighoring pastoralists cells.

While I cannot fully reject that other unobserved shocks simultaneously determine conflict and migration, conflicts at origin are controlled for in all specifications within a range of 50 km, and the results are robust to an alternative buffer of 25km as shown in Tables A16 and A17 in appendix. Instead, average conflict at destination and on migration routes exploits variation in conflicts located further away from the locality, similar to the shift-share instruments used in the migration and trade literature, with an additional layer of variation for route conflicts depending on the spatial location of localities and destination. To ensure that destination shocks are exogenous to the locality conditions, I exclude the own region from the set of possible destinations. Further robustness checks should exclude all conflicts located within 25 km around a locality from the average route insecurity measure.

Concerns of reverse causality are alleviated by the precise temporal definition of the migration and conflict variables.

Misspecification bias and measurement error I address potential mesurement issues by showing the main gravity and aggregate results are robust to an alternative buffer of 5km instead of 2km around the shortest road between origin and destination in Tables A7 to A9, to the inclusion of less precisely geo-coded events in Tables A10 to A12 and to using the count of fatalities instead of the count of conflicts in Tables A13 to A15 . The geographical correlation between past conflict occurence and perceptions of insecurity in Table 13 further alleviates concerns about the quality of ACLED data geo-coding.

External validity While the continuation of data collection efforts by the Malian National Institute of Statistics (INSTAT) in such a instable context is impressive and provides us with a unique dataset to study the impacts of conflict, the representativity of the household survey cannot be guaranteed and the hundreds of thousands of Malians displaced by the conflict either internally or in neighboring countries are missing from the survey data. I match the exhaustive list of rural localities of Mali from the 2009 population census with the successive round of EMOP household data to study how conflict intensity affected the probability of localities to be included in the survey sample. The results of this panel analysis displayed in Table A4 in appendix show that the probability of census locality to be included in the EMOP survey sample is much lower for localities surrounded by conflict events in the current and past year, but is not affected by conflicts happening around the shortest road between the locality and Bamako, where the headquarter of the INSTAT is located. This implies that the estimated effect of route insecurity on migration may not be representative for localities directly affected by conflict.

Mechanisms: perceived insecurity and liquidity constraints Insecurity on migration routes can deter mobility of household members either directly to avoid exposure to conflict or indirectly by increasing migration costs.

The avoidance mechanism is more credible if households have information on such conflict incidents occurring further from their homes along the route to usual destinations. To test for the plausibility of an information channel, I assess the impact of conflict events on the perception of insecurity. The strong correlation between conflict fatalities around localities and perceived insecurity presented in Table 13 confirms that households are informed about conflicts occuring far from their locality and confirms the main analysis' premise that past conflict affect perception of current insecurity. While consistent with lower information about conflicts located further away from a given locality, the decaying effect of conflicts with distance to the locality could also signal that households naturally feel less threatened by more distant conflicts.

The liquidity constraint mechanism whereby migration is restricted by the unability to finance migration costs is well documented in the migration litterature, in particular in association with weather shocks (Bryan et al., 2014; Jayachandran, 2006; Kleemans, 2015). In the case of Mali, Focus Group Discussions I conducted in Bamako in 2020 with two groups of young female migrants from several regions of the country, all women reported traveling by bus, and that either their parents or husband financed the cost of the travel. According to a 2022 article from the Malian media Maliweb, [10] prices of tickets to Bamako ranged between 3000 FCFA from Segou, in the center of the country to 35000 FCFA from Gao in the north of the country, compared to an average

[10] https://www.maliweb.net/economie/transport/cout-du-transport-routier-au-mali-les-transporteurs-augm html

monthly wage of 15000 FCFA in the survey sample. Qualitative interviews I conducted in 2022 via phone with commercial agents from transport companies operating in Mali suggest these costs may increase in dangerous routes due to re-routing onto alternative itineraries. [11]

The negative impact of route conflict on post-harvest departures for localities affected by drougths estimated in Table 9, and driven by the poorest households in the sample as evidence by subsample analysis in Table 11, suggest liquidity constraints may become binding for households affected by multiple shocks: lower agricultural income due to harvest losses and higher migration costs due to route insecurity. On the other hand, the stronger average effect of route insecurity for the richest households, women and young men suggests that these groups may exhibit a greater inclination to avoid insecure migration.

Thus, the findings align with the notion that both avoidance and liquidity constraints likely influence migration outcomes in a setting with insecurity on migration routes.

Table 13: Conflicts and perceived insecurity

	Tensions in community (dummy)			Victim future (dummy)			Feels unsafe count		
	Maximum distance of conflicts to village								
	15 km	*25 km*	*50 km*	*15 km*	*25 km*	*50 km*	*15 km*	*25 km*	*50 km*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
All fatalities (ihs)	0.171***	0.093*	0.091*	0.230***	0.169***	0.122***	0.112***	0.081***	0.066***
	(0.060)	(0.054)	(0.046)	(0.073)	(0.051)	(0.042)	(0.026)	(0.025)	(0.024)
Mean DV	0.277	0.277	0.277	0.245	0.245	0.245	1.169	1.169	1.169
Standard Deviation	0.448	0.448	0.448	0.430	0.430	0.430	1.643	1.643	1.643
Observation	13416	13416	13416	23585	23585	23585	26232	26232	26232
Localities	406	406	406	689	689	689	789	789	789

Notes : Estimation with PPML. Robust standard errors in parentheses clustered at the locality level - *** p<0.01, ** p<0.05, * p<0.1. The sample period includes all years 2011 and 2013 to 2019. Conflict variables are defined in terms of count of fatalities and measured between the first and second quarter of data collection.

Welfare implications: road insecurity, migration and trade If we assimilate road conflict as a shock to migration market access, these conflicts are also likely to affect market access for producers and consumers. Morten and Oliveira (2018) finds that 76% or welfare gains from road construction in Brazil can be attributed to reduced trade costs, while only 34% can be attributed to reduced migration costs. Hence, further analysis estimating the relative elasticity of trade and migration to conflict is required to asses the share of consumption losses that can be attributed to reduced migration.

[11] Namely Toupac, Tilemsi, and Nour Transport. Notably, it was revealed that certain routes, such as the direct connection between Bamako and Tombouctou offered by Nour Transport, had been disrupted due to security concerns. Alternative routes through Niger and Burkina Faso significantly extend the journey by 2.5 times its original distance, which would also significantly increase the cost of travel. Furthermore, responses from a commercial respondent associated with Tilemsi indicated that boarding between official stops was no longer feasible as it was considered too risky. This could also further increase travel costs by necessitating additional stretches to reach bus terminals.

7 Conclusion

This paper documents a previously unstudied barrier to migration flows in conflict-affected areas: insecurity on migration routes. Estimating a gravity model of migration, including route conflicts as an additional bilateral migration cost, I find that economic migration and reception of remittances are reduced for migration corridors affected by conflict, with a stronger effect in localities affected by climatic shocks. The gravity estimates indicate large substitution between destinations for these economic migrants, and I can detect no significant average effect of average route insecurity on economic migration in the reduced form analysis.

Such substitution does not fully offset the effect of route conflicts on other types of mobility. Combining high-frequency data on post-harvest departures of household members and conflict incidence, I find that conflicts between rural localities and pre-conflict migration destinations reduce the effect of droughts on the probability of temporary departures in the post-harvest season. This constrained mobility bears significant welfare implications, as evidenced by the varied migratory response to droughts, encompassing remittances from economic migrants, departures of adult household members for permanent migration, and child fostering. The reduced form estimates show a stark reduction in consumption per capita of households in localities which faced high average insecurity on their usual migration routes shortly before the seasonal peak of migration departures, and increase in food insecurity in localities additionally hit by droughts. The reduced-form results for post-harvest departures suggest that this may be the consequence of reduced flexibility in household composition.

The findings imply that households facing multiple shocks may struggle to cope simultaneously: poorer households may have lower capacity to mitigate their exposure to conflict by avoiding risky migration routes, and when confronted with both conflict and climatic shocks, these may reinforce each other, rendering migration financially untenable. Moreover, the results suggest that even populations not directly affected by conflict may experience heightened vulnerability due to the diminished accessibility of their usual migration destinations.

References

Aggarwal, Shilpa (2018). "Do rural roads create pathways out of poverty? Evidence from India". In: *Journal of Development Economics*. ISSN: 03043878. DOI: 10.1016/j.jdeveco.2018.01.004.

Akresh, Richard (2009). "Risk, Network Quality, and Family Structure: Child: Child Fostering Decisions in Burkina Faso". In: *Journal of Human Resources*.

Arias, María Alejandra, Ana María Ibáñez, and Andrés Zambrano (July 2019). "Agricultural production amid conflict: Separating the effects of conflict into shocks and uncertainty". In: *World Development* 119, pp. 165–184. ISSN: 18735991. DOI: 10.1016/j.worlddev.2017.11.011.

Azam, Jean-Paul and Flore Gubert (2005). "Those in kayes the impact of remittances on their recipients in Africa". In: *Revue Economique*. ISSN: 00352764.

Beine, Michel, Simone Bertoli, and Jesús Fernández-Huertas Moraga (2016). "A Practitioners' Guide to Gravity Models of International Migration". In: *World Economy* 39(4), pp. 496–512. ISSN: 14679701. DOI: 10.1111/twec.12265.

Beine, Michel, Frédéric Docquier, and Çağlar Özden (2011). "Diasporas". In: *Journal of Development Economics* 95(1), pp. 30–41. ISSN: 03043878. DOI: 10.1016/j.jdeveco.2009.11.004.

Beine, Michel, Christopher Parsons, Source The, Scandinavian Journal, No April, Michel Beine, and Christopher Parsons (2021). "Climatic Factors as Determinants of International Migration Published by : Wiley on behalf of The Scandinavian Journal of Economics Climatic Factors as Determinants of International Migration *". In: 117(2), pp. 723–767.

Bertoli, Simone and Jesús Fernández-Huertas Moraga (2013). "Multilateral resistance to migration". In: *Journal of Development Economics* 102(5958), pp. 79–100. ISSN: 03043878. DOI: 10.1016/j.jdeveco.2012.12.001.

Blattman, Christopher and Edward Miguel (2010). "Civil war ". In: *Journal of Economic Literature*.

Blumenstock, Joshua E. (2012). "Information Technology for Development Inferring patterns of internal migration from mobile phone call records : evidence from Rwanda". In: *Information Technology for Development*.

Boudinaud, Laure and Sacha Alex Orenstein (2021). "ASSESSING CROPLAND ABANDONMENT from VIOLENT CONFLICT in CENTRAL MALI with SENTINEL-2 and GOOGLE EARTH ENGINE". In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. DOI: 10.5194/isprs-archives-XLVI-4-W2-2021-9-2021.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh". In: *Econometrica* 82(5), pp. 1671–1748. ISSN: 0012-9682. DOI: 10.3982/ecta10489.

Büttner, Nicolas, Michael Grimm, and Sidiki Soubeiga (June 2021). "Political instability and households' investment behavior: Evidence from Burkina Faso". In: *Journal of Comparative Economics* 50(2), pp. 350–368. ISSN: 10957227. DOI: 10.1016/j.jce.2021.11.003.

Calvo, Thomas, Emmanuelle Lavallée, Mireille Razafindrakoto, and François Roubaud (2020). "Fear Not For Man? Armed conflict and social capital in Mali". In: *Journal of Compar-*

ative Economics. ISSN: 10957227. DOI: 10.1016/j.jce.2019.11.005. URL: https://hal.archives-ouvertes.fr/hal-02446972.
Chort, Isabelle and Maëlys de la Rupelle (2016). "Determinants of Mexico-U.S. Outward and Return Migration Flows: A State-Level Panel Data Analysis". In: *Demography* 53(5). ISSN: 15337790. DOI: 10.1007/s13524-016-0503-9.
Clemens, Michael Andrew (2020). "Migration from Developing Countries: Selection, Income Elasticity and Simpson's Paradox". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3688065.
Di Maio, Michele, Valerio Leone Sciabolazza, and Vasco Molini (Mar. 2023). "Migration in Libya: A spatial network analysis". In: *World Development* 163, p. 106139. ISSN: 0305-750X. DOI: 10.1016/J.WORLDDEV.2022.106139.
Donaldson, Dave (2018). "Railroads of the Raj: Estimating the impact of transportation infrastructure". In: *American Economic Review* 108(4-5). ISSN: 00028282. DOI: 10.1257/aer.20101199.
Donaldson, Dave and Richard Hornbeck (2016). "Railroads and American economic growth: A "market access" approach". In: *Quarterly Journal of Economics* 131(2). ISSN: 15314650. DOI: 10.1093/qje/qjw002.
Dube, Oeindrila and Juan F. Vargas (Oct. 2013). "Commodity price shocks and civil conflict: Evidence from Colombia". In: *Review of Economic Studies* 80(4), pp. 1384–1421. ISSN: 00346527. DOI: 10.1093/restud/rdt009.
Gray, C. L. and V. Mueller (2012). "Natural disasters and population mobility in Bangladesh". In: *Proceedings of the National Academy of Sciences*. ISSN: 0027-8424.
Harari, Mariaflavia and Eliana La Ferrara (2018). "Conflict, climate, and cells: A disaggregated analysis". In: *Review of Economics and Statistics*. ISSN: 15309142. DOI: 10.1162/rest{_}a{_}00730.
Jayachandran, Seema (2006). "Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries". In: *Journal of Political Economy*. ISSN: 0022-3808. DOI: 10.1086/503579.
Kleemans, Marieke (2015). "Migration Choice under Risk and Liquidity Constraints". In: *Working Paper*.
Libois, François (2016). *Households in Times of War : Adaptation Strategies during the Nepal Civil War*. Tech. rep.
Mayda, Anna Maria (2010). "International migration: A panel data analysis of the determinants of bilateral flows". In: *Journal of Population Economics* 23(4), pp. 1249–1274. ISSN: 09331433. DOI: 10.1007/s00148-009-0251-x.
McGuirk, Eoin and Nathan Nunn (2021). "Transhumant Pastoralism, Climate Change, and Conflict in Africa". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.3753152.
Miguel, Edward, Shanker Satyanath, and Ernest Sergenti (2004). "Economic shocks and civil conflict: An instrumental variables approach". In: *Journal of Political Economy*. ISSN: 00223808. DOI: 10.1086/421174.
Morten, Melanie and Jaqueline Oliveira (2018). "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City". In: *NBER Working Paper*.
Ortega, Francesc and Giovanni Peri (2013). "The effect of income and immigration policies on international migration". In: *Migration Studies* 1(1). ISSN: 2049-5838. DOI: 10.1093/migration/mns004.

Ortega, Francesc and Giovanni Peri (2021a). "The Role of Income and Immigration Policies in Attracting International Migrants". In: *SSRN Electronic Journal*(6655). DOI: 10.2139/ssrn.2089704.

Ortega, Francesc and Giovanni Peri (2021b). "The Role of Income and Immigration Policies in Attracting International Migrants". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2089704.

Rockmore, Marc (Oct. 2017). "The cost of fear: The welfare effect of the risk of violence in Northern Uganda". In: *World Bank Economic Review* 31(3), pp. 650–669. ISSN: 1564698X. DOI: 10.1093/wber/lhw025.

Rohner, Dominic and Mathias Thoenig (2021). *The Elusive Peace Dividend of Development Policy: From War Traps to Macro Complementarities*. DOI: 10.1146/annurev-economics-073120-102652.

Shrestha, Maheshwor (2017). *Push and Pull: A Study of International Migration from Nepal*. DOI: 10.1596/1813-9450-7965.

Tapsoba, Augustin (2023). *The Cost of Fear: Impact of Violence Risk on Child Health During Conflict*.

Verpoorten, Marijke (2009). "Household coping in war- and peacetime: Cattle sales in Rwanda, 1991-2001". In: *Journal of Development Economics*. ISSN: 03043878. DOI: 10.1016/j.jdeveco.2008.01.003.

A Appendix

A.1 Characteristics of economic migrants

Although the general economic migration module does not contain information on the characteristics of migrants, these characteristics can be inferred from the 2016 survey which includes an additional detailed migration module. This additional module is based on different inclusion criteria of migrants: while it is not restricted to working migrants as the general economic migration module used in this study, only migrants who left at least six months ago are recorded. Table A1 in appendix shows that while both definitions yield about the same count of reported migrants, only roughly two thirds of these are overlapping in terms of households reporting migrants in the same destination. This implies that the six months inclusion criteria is binding for about a third of migrants, with high heterogeneity between internal and international migrants: while only 63% of internal migrants reported in the economic migration module are also declared in the detailed module, almost all economic migrants in ECOWAS countries are reported in the detailed module (97% of them). This is consistent with longer average length of migration episodes reported by migrants abroad in the detailed module, as displayed in table **??** and **??**.

Those migrants with detailed data are a selected subsample of economic migrants having left their household of origin more than 6 months ago. Although there is no strong reason ex-ante to expect these recent migrants would differ in any other characteristics than the average length of their migration, I inpute lower and upper bound values for the more recent migrants [12] to compute upper and lower bounds of average characteristics of average migrants. If we believe the definition of the detailed migration module, then all ommitted economic migrants had left less than 6 months ago, which implies that the total share of internal economic migrants having left less than a year ago if about 39%, while it represents only about 26% of economic migrants to ECOWAS countries. While I cannot exclude share of economic migrants in urban destinations as low as 54% for internal migration and 65% for ECOWAS countries in case all recent economic migrants unobserved in the detailed module moved to rural areas, the reality is probably closer to the observed average of respectively 70% and 67% of all economic migrants in Mali and ECOWAS countries.

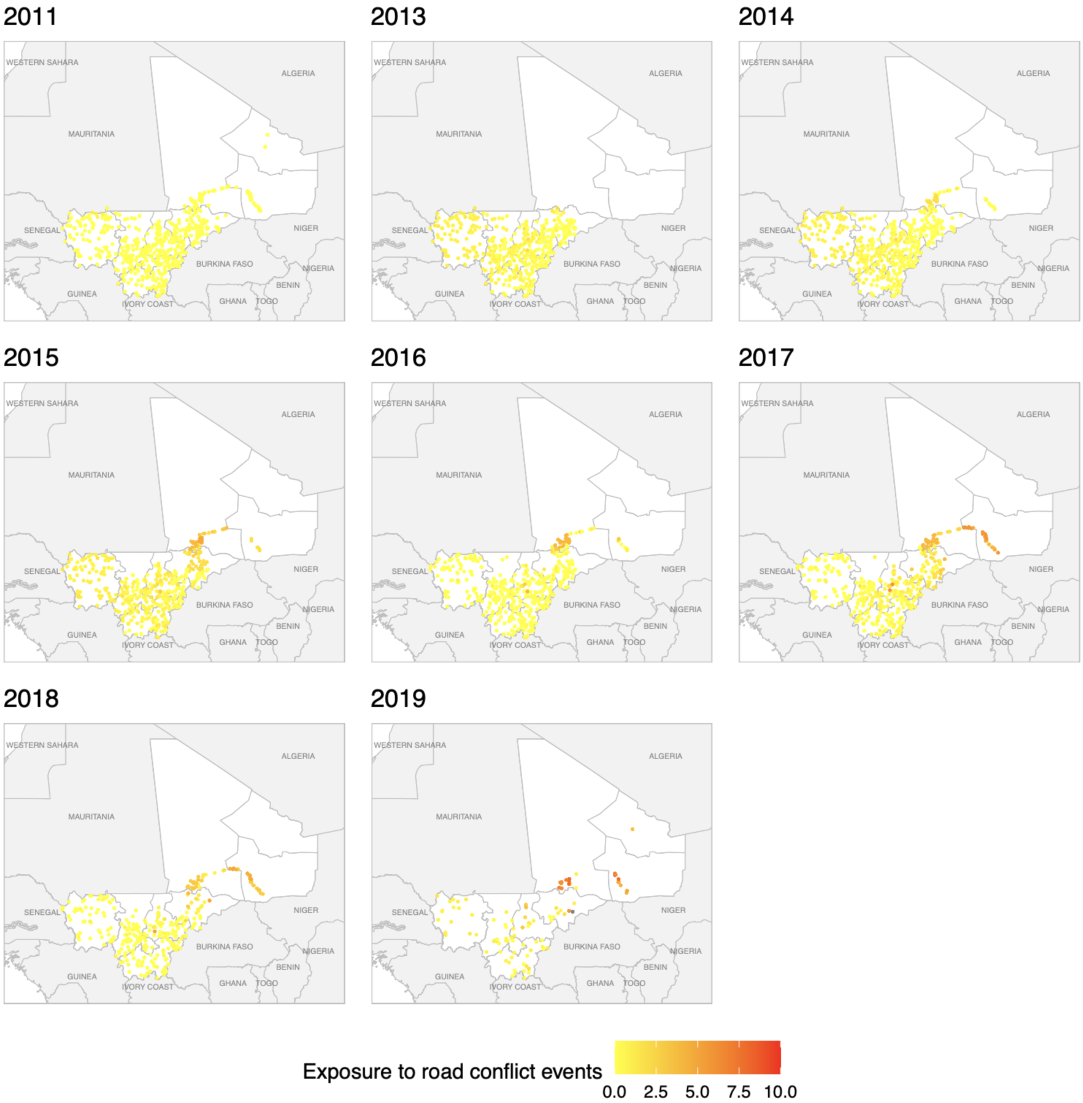
Table A1: Count of migrants according to data sources

	(1) Detailed module (departure $\geq$ 6mo)	(2) General module (migrant working)	(3) Intersection
All destinations	1535	1461	921
Mali	905	843	651
ECOWAS	320	278	270
Men	1397	1328	844
Women	147	133	77

[12] I inpute 0 or 1 as lower and upper bound for dummy variables, and the 25th and 75th percentile value for continuous variables

A.2 Supplementary figures and tables

Figure A 1: Average exposure to conflict on migration routes



Notes: Rural localities included in the study sample only. The average exposure is the average number of conflicts events within 2km of the shortest route to the main city of each region of Mali, excluding the region of the locality, weighted by the share of return migrants from each of these regions as measured in the 2009 population census.

Table A2: Average route insecurity and post-harvest departures by age groups

	Men						Women					
	All departures		Temporary		Permanent		All departures		Temporary		Permanent	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: 0-14												
Route conflicts (ihs)	-0.366	-0.187	-0.604*	-0.412*	0.269	0.437	-0.420***	-0.384*	-0.543***	-0.548**	-0.050	0.083
	(0.376)	(0.315)	(0.326)	(0.245)	(0.489)	(0.497)	(0.146)	(0.211)	(0.194)	(0.267)	(0.175)	(0.224)
Orig. conflict (ihs)	0.240	0.210	0.454***	0.416***	-0.556	-0.563	0.078	0.072	0.125	0.126	0.004	-0.025
	(0.232)	(0.230)	(0.152)	(0.148)	(0.558)	(0.573)	(0.176)	(0.182)	(0.167)	(0.172)	(0.282)	(0.290)
Dest. conflict (ihs)	-0.334***	-0.342***	-0.239**	-0.256***	-0.867***	-0.827***	-0.012	-0.014	0.117	0.118	-0.434	-0.436
	(0.111)	(0.108)	(0.097)	(0.093)	(0.287)	(0.288)	(0.100)	(0.103)	(0.165)	(0.169)	(0.324)	(0.332)
Orig SPEI drought frequency	0.275***	0.333***	0.230*	0.289**	0.479***	0.551***	0.134	0.146	0.189	0.188	-0.021	0.047
	(0.091)	(0.095)	(0.128)	(0.140)	(0.106)	(0.132)	(0.117)	(0.137)	(0.124)	(0.144)	(0.145)	(0.158)
Route conflicts (ihs) x Orig SPEI drought frequency		-0.455***		-0.499***		-0.417		-0.074		0.010		-0.278
		(0.119)		(0.136)		(0.303)		(0.158)		(0.218)		(0.219)
Observations	34447	34447	30302	30302	14177	14177	32212	32212	28935	28935	13188	13188
Pseudo R2	0.08	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.06	0.06
Panel B: 15-24												
Route conflicts (ihs)	0.006	0.052	0.076	0.114	-0.140	0.021	-0.167	-0.144	-0.020	0.096	-0.620**	-0.974**
	(0.116)	(0.153)	(0.208)	(0.218)	(0.193)	(0.226)	(0.113)	(0.143)	(0.123)	(0.165)	(0.311)	(0.456)
Orig. conflict (ihs)	0.100	0.090	0.050	0.043	0.242	0.204	0.264	0.262	0.241	0.229	0.325	0.384
	(0.092)	(0.096)	(0.105)	(0.104)	(0.242)	(0.244)	(0.208)	(0.210)	(0.204)	(0.204)	(0.326)	(0.348)
Dest. conflict (ihs)	-0.114	-0.118	-0.029	-0.033	-0.331*	-0.334**	0.107	0.106	0.251***	0.240**	-0.404	-0.441*
	(0.081)	(0.078)	(0.098)	(0.097)	(0.172)	(0.159)	(0.088)	(0.087)	(0.097)	(0.096)	(0.246)	(0.245)
Orig SPEI drought frequency	0.063	0.082	-0.024	-0.010	0.280*	0.373*	0.220**	0.228***	0.177**	0.213**	0.439***	0.273
	(0.061)	(0.067)	(0.093)	(0.103)	(0.143)	(0.199)	(0.089)	(0.088)	(0.089)	(0.093)	(0.147)	(0.166)
Route conflicts (ihs) x Orig SPEI drought frequency		-0.095		-0.080		-0.304		-0.049		-0.247*		0.693**
		(0.138)		(0.092)		(0.390)		(0.120)		(0.147)		(0.326)
Observations	11542	11542	10345	10345	6314	6314	12296	12296	11310	11310	4740	4740
Pseudo R2	0.09	0.09	0.09	0.09	0.12	0.12	0.09	0.09	0.09	0.09	0.11	0.11
Panel C: 24-64												
Route conflicts (ihs)	-0.200	-0.116	-0.197	0.040	-0.184	-0.182	-0.404	-0.325	-0.495	-0.265	-0.130	-0.070
	(0.160)	(0.172)	(0.194)	(0.144)	(0.246)	(0.318)	(0.258)	(0.313)	(0.368)	(0.373)	(0.363)	(0.446)
Orig. conflict (ihs)	0.234**	0.225**	0.220*	0.192*	0.370	0.370	0.088	0.073	-0.160	-0.206	0.657	0.643
	(0.111)	(0.108)	(0.120)	(0.107)	(0.276)	(0.277)	(0.218)	(0.222)	(0.285)	(0.272)	(0.424)	(0.427)
Dest. conflict (ihs)	-0.104	-0.111	-0.040	-0.073	-0.212	-0.212	-0.164	-0.173	-0.042	-0.070	-0.480	-0.486
	(0.078)	(0.080)	(0.153)	(0.157)	(0.146)	(0.146)	(0.122)	(0.123)	(0.195)	(0.198)	(0.329)	(0.324)
Orig SPEI drought frequency	0.161	0.190	-0.060	0.023	0.515***	0.516***	0.219	0.240	0.018	0.098	0.465***	0.477***
	(0.149)	(0.146)	(0.147)	(0.165)	(0.149)	(0.129)	(0.143)	(0.150)	(0.107)	(0.134)	(0.158)	(0.158)
Route conflicts (ihs) x Orig SPEI drought frequency		-0.166		-0.645***		-0.005		-0.136		-0.725**		-0.066
		(0.117)		(0.116)		(0.210)		(0.201)		(0.344)		(0.244)
Observations	19738	19738	16222	16222	12238	12238	19799	19799	16747	16747	9572	9572
Pseudo R2	0.07	0.07	0.08	0.08	0.09	0.09	0.07	0.07	0.06	0.06	0.11	0.11

Notes :Estimation with PPML. Robust standard errors clustered by year and locality are reported in parentheses. All regressions include year and locality fixed effects. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

A.2.1 Determinants of conflicts

Table A3: Weather shocks and conflict: grid analysis

	Dependent variable:							
	No lags		Temporal lag		Spatial Durbin		Spatial lags in X and Y	
	event_count	fatalities_count	event_count	fatalities_count	event_count	fatalities_count	event_count	fatalities_count
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
drought_t	0.162**	0.498*	0.076	0.386	−0.027	−0.142	0.001	−0.085
	(0.066)	(0.268)	(0.053)	(0.253)	(0.113)	(0.462)	(0.077)	(0.390)
drought_{t-1}	0.150**	0.712***	0.045	0.534**	−0.022	−0.159	0.018	−0.093
	(0.062)	(0.254)	(0.051)	(0.240)	(0.112)	(0.457)	(0.077)	(0.386)
event count_{t-1}			0.775***				0.366***	
			(0.017)				(0.018)	
$\text{fatalities count}_{t-1}$				0.405***				0.147***
				(0.019)				(0.019)
$\text{event count spatial lag}_t$							0.787***	
							(0.030)	
$\text{event count spatial lag}_{t-1}$							0.062	
							(0.044)	
$\text{fatalities count spatial lag}_t$								0.704***
								(0.035)
$\text{fatalities count spatial lag}_{t-1}$								0.334***
								(0.046)
$\text{drought spatial lag}_t$					0.273*	0.889	−0.010	0.201
					(0.142)	(0.577)	(0.097)	(0.487)
$\text{drought spatial lag}_{t-1}$					0.233*	1.220**	−0.053	0.173
					(0.137)	(0.559)	(0.094)	(0.473)
Observations	4,250	4,250	4,250	4,250	4,250	4,250	4,250	4,250
R^2	0.431	0.274	0.625	0.352	0.432	0.275	0.736	0.484
Adjusted R^2	0.367	0.191	0.583	0.278	0.367	0.192	0.705	0.425

Note: Each observation corresponds to a 0.5° cell-year."Drought" is defined as the number of months during the current year growing season for which the Standardadized Potential Evapo-Transpiration Index is below -1.5. Spatial lags includes up to eight contiguous cells. Standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

A.2.2 Determinants of sample inclusion

Table A4: Conflict and sample inclusion of census localities

	All census localities	Localities surveyed at least once					
		Locality surveyed					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	-2.69***						
	(0.035)						
Conflicts route to Bamako (ihs)	0.013	0.036		0.017		0.013	
	(0.030)	(0.127)		(0.127)		(0.126)	
Orig. conflicts, 50 km (ihs)	0.016	-0.097***					
	(0.026)	(0.031)					
Conflicts route to Bamako (ihs), t-1			0.051		0.041		0.032
			(0.095)		(0.101)		(0.103)
Orig. conflicts, 50 km (ihs), t-1			-0.120***				
			(0.037)				
Orig. conflicts, 25 km (ihs)				-0.092**			
				(0.036)			
Orig. conflicts, 25 km (ihs), t-1					-0.124***		
					(0.035)		
Orig. conflicts, 15 km (ihs)						-0.110**	
						(0.043)	
Orig. conflicts, 15 km (ihs), t-1							-0.109**
							(0.052)
Pseudo R^2	3.39×10^{-5}	0.1169	0.1173	0.1167	0.1170	0.1167	0.1168
Observations	81,352	13,600	13,600	13,600	13,600	13,600	13,600
Dependent variable mean	0.0685	0.4097	0.4097	0.4097	0.4097	0.4097	0.4097
Locality and Year fixed effects		✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The sample period includes all years 2011 and 2013 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest.

A.3 Robustness checks

A.3.1 Robustness of gravity model to IV estimation

Table A5: Robustness check road conflicts, internal migration and remittances: instrumental variable

Dependent Variables:	Road conflicts	# internal migrants		Any transfer		Amount (FCFA)	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Variables							
Road conflicts neighboring cell, tm1	30.22***						
	(3.486)						
idistancert		-0.4103***	-0.3244***	-0.3537***	-0.2210***	-0.2799***	-0.0646
		(0.0445)	(0.0860)	(0.0353)	(0.0600)	(0.0668)	(0.1559)
Road conflicts		-0.2256***	-0.4657**	-0.1278**	-0.3967***	-0.0859*	-0.5210*
		(0.0405)	(0.1952)	(0.0503)	(0.1079)	(0.0439)	(0.2831)
Origin conflicts		-0.0454	0.0012	-0.0118	0.0315	-0.0289	0.0375
		(0.0537)	(0.0548)	(0.0547)	(0.0417)	(0.0575)	(0.0625)
Dest conflicts		-0.0917***	-0.0156	-0.0302	0.0207	0.0816	0.1615***
		(0.0290)	(0.0639)	(0.0295)	(0.0286)	(0.0696)	(0.0601)
resid_1st_stage			0.2476		0.2760**		0.4406
			(0.2022)		(0.1085)		(0.2852)
Fixed-effects							
year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LOCALITE09	Yes	Yes	Yes	Yes	Yes	Yes	Yes
destcode	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	42,388	22,936	22,936	39,540	39,492	12,724	12,712
Squared Correlation	0.55415	0.20332	0.20178	0.17593	0.17668	0.19416	0.19710
Pseudo R^2	0.99997	0.25088	0.25121	0.25690	0.25721	1.1783	1.1781

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

A.3.2 Robustness of gravity model to inclusion of pair fixed effects

Table A6: Gravity model of internal migration: robustness with pair fixed-effects

	Internal migrants/Stayers				
	(1)	(2)	(3)	(4)	(5)
Route conflicts (ihs), t-1	-0.138	-0.226**	-0.218**	-0.286***	-0.239*
	(0.122)	(0.097)	(0.100)	(0.070)	(0.134)
Orig. conflicts (ihs), t-1	-0.056		-0.040		
	(0.058)		(0.108)		
Dest. conflicts (ihs), t-1	-0.144		0.008		
	(0.105)		(0.061)		
Pseudo R^2	0.1867	0.2258	0.2258	0.2623	0.2314
Observations	2,658	2,658	2,658	1,879	2,614
Dependent variable mean	0.0382	0.0382	0.0382	0.0540	0.0388
Implied effect for one additional conflict	-0.006	-0.010	-0.010	-0.023	-0.012
Year fixed effects	✓	✓	✓	✓	✓
Destination fixed effects	✓		✓	✓	
Locality fixed effects	✓		✓		✓
Pair fixed effects		✓	✓	✓	✓
Locality-Year fixed effects				✓	
Destination-Year fixed effects					✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The p-values displayed at the bottom of the table are only for the route conflict coefficients, and conley standard errors allow fo spatial correlations up to 50km. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

A.3.3 Robustness to larger 5km buffers around shortest road

Table A7: Gravity model of internal migration and remittances by drought at origin: robustness to 5km buffer for route conflicts

	Internal migrants/Stayers		Any transfer		Amount received (FCFA)	
	(1)	(2)	(3)	(4)	(5)	(6)
Route conflicts (ihs), t-1	-1.32***	-1.12***	-0.663***	-0.630***	-0.606***	-0.388***
	(0.223)	(0.207)	(0.159)	(0.153)	(0.223)	(0.115)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency (ihs)		-0.507***		-0.140		-0.906***
		(0.191)		(0.288)		(0.234)
Pseudo R^2	0.3097	0.3108	0.3289	0.3291	1.060	1.058
Observations	15,294	15,294	13,651	13,651	3,624	3,624
Dependent variable mean	0.0182	0.0182	0.0373	0.0373	4,787.7	4,787.7
Locality-Year fixed effects	✓	✓	✓	✓	✓	✓
Destination-Year fixed effects	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and pair are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflicts includes all fatalities within a 50 km buffer of the locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

Table A8: Average route insecurity and economic migration: robustness to 5km buffer for route conflicts

	Economic migrants in Mali									
	Internal migrants/Stayers		Any transfer		Amount (FCFA)		Av. dest. wage		Av. dest. dist.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts within 5km (ihs), t-1	0.043	-0.034	0.010	0.012	0.047	0.027	-0.015	-0.015	-0.014	-0.009
	(0.168)	(0.192)	(0.099)	(0.106)	(0.118)	(0.131)	(0.022)	(0.022)	(0.049)	(0.052)
Orig. conflicts (ihs), t-1	0.003	0.026	-0.027	-0.028	-0.030	-0.023	-0.043*	-0.043*	0.038	0.037
	(0.131)	(0.130)	(0.099)	(0.104)	(0.096)	(0.100)	(0.024)	(0.024)	(0.050)	(0.051)
Orig. SPEI drought frequency, t-1	0.037	-0.150	0.328***	0.335***	0.301***	0.254**	-0.010	-0.009	-0.074**	-0.064*
	(0.114)	(0.179)	(0.082)	(0.128)	(0.071)	(0.100)	(0.013)	(0.016)	(0.032)	(0.034)
Dest. conflicts (ihs), t-1	0.052	0.072	-0.089	-0.090	-0.138*	-0.132	-0.039**	-0.039**	0.0008	-0.0004
	(0.110)	(0.111)	(0.057)	(0.057)	(0.078)	(0.081)	(0.017)	(0.017)	(0.054)	(0.055)
Route conflicts within 5km (ihs), t-1 × Orig. SPEI drought frequency, t-1		0.280*		-0.012		0.081		-0.002		-0.016
		(0.151)		(0.113)		(0.110)		(0.017)		(0.037)
Pseudo R^2	0.1464	0.1468	0.1487	0.1487	1.211	1.211	1.001	1.001	1.013	1.013
Observations	14,901	14,901	16,956	16,956	16,323	16,323	10,675	10,675	10,675	10,675
Dependent variable mean	0.0847	0.0847	0.0155	0.0155	8,046.3	8,046.3	37,231.1	37,231.1	451.2	451.2
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

Table A9: Average route insecurity and households mobility: robustness to 5km buffer for route conflicts

	Post-harvest departures								**Early season returns**	
	All departures		Temporary		Permanent		Attrition		Return mig. in q1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts within 5km (ihs)	-0.293**	-0.213	-0.290*	-0.158	-0.286	-0.228	0.391	0.345	-0.114	-0.045
	(0.143)	(0.143)	(0.155)	(0.124)	(0.215)	(0.269)	(0.309)	(0.419)	(0.102)	(0.160)
Orig. conflicts (ihs)	0.318*	0.301	0.318*	0.289	0.385	0.375	1.62**	1.64**	-0.372*	-0.388*
	(0.186)	(0.189)	(0.182)	(0.186)	(0.347)	(0.348)	(0.795)	(0.798)	(0.205)	(0.203)
Dest. conflicts (ihs)	-0.028	-0.036	0.085	0.067	-0.343**	-0.344**	-0.347	-0.347	-0.042	-0.050
	(0.054)	(0.054)	(0.086)	(0.083)	(0.136)	(0.138)	(0.571)	(0.578)	(0.089)	(0.093)
Orig SPEI drought frequency	0.165**	0.197**	0.087	0.138	0.395***	0.422***	0.176	0.126	0.086	0.107
	(0.081)	(0.083)	(0.086)	(0.098)	(0.086)	(0.079)	(0.238)	(0.195)	(0.087)	(0.090)
Route conflicts within 5km (ihs) × Orig SPEI drought frequency		-0.160**		-0.304***		-0.096		0.133		-0.108
		(0.076)		(0.086)		(0.177)		(0.512)		(0.135)
Pseudo R^2	0.0669	0.0670	0.0743	0.0748	0.0913	0.0914	0.2237	0.2238	0.1055	0.1056
Observations	180,709	180,709	176,351	176,351	134,758	134,758	4,776	4,776	169,460	169,460
Dependent variable mean	0.0518	0.0518	0.0388	0.0388	0.0186	0.0186	0.0773	0.0773	0.0536	0.0536
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

A.3.4 Robustness to inclusion of less precisely geo-coded events

Table A10: Gravity model of internal migration and remittances by drought at origin: robustness to less precisely geocoded events

	Internal migrants/Stayers		Any transfer		Amount received (FCFA)	
	(1)	(2)	(3)	(4)	(5)	(6)
Route conflicts (ihs), t-1	-0.876***	-0.676***	-0.443***	-0.448***	-0.549***	-0.422***
	(0.214)	(0.217)	(0.086)	(0.090)	(0.141)	(0.112)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency (ihs)		-0.435***		0.018		-0.335*
		(0.150)		(0.169)		(0.190)
Pseudo R^2	0.3060	0.3071	0.3275	0.3275	1.060	1.059
Observations	15,294	15,294	13,651	13,651	3,624	3,624
Dependent variable mean	0.0182	0.0182	0.0373	0.0373	4,787.7	4,787.7
Locality-Year fixed effects	✓	✓	✓	✓	✓	✓
Destination-Year fixed effects	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and pair are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflicts includes all fatalities within a 50 km buffer of the locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

Table A11: Average route insecurity and economic migration: robustness to less precisely geocoded events

	Economic migrants in Mali									
	Internal migrants/Stayers		Any transfer		Amount (FCFA)		Av. dest. wage		Av. dest. dist.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs), t-1	-0.201	-0.296*	-0.100	-0.109	-0.096	-0.127	-0.013	-0.009	-0.050	-0.043
	(0.126)	(0.152)	(0.117)	(0.126)	(0.095)	(0.112)	(0.025)	(0.027)	(0.042)	(0.045)
Orig. conflict (ihs), t-1	-0.047	-0.016	0.001	0.006	0.005	0.018	-0.036	-0.037	0.020	0.017
	(0.106)	(0.106)	(0.079)	(0.082)	(0.080)	(0.085)	(0.026)	(0.026)	(0.043)	(0.044)
Orig. SPEI drought frequency (ihs)	0.030	-0.094	0.317***	0.300***	0.291***	0.239***	-0.011	-0.006	-0.074**	-0.063*
	(0.105)	(0.121)	(0.082)	(0.105)	(0.070)	(0.090)	(0.012)	(0.013)	(0.029)	(0.034)
Dest. conflict (ihs), t-1	0.042	0.062	-0.084	-0.082	-0.100	-0.089	-0.025	-0.025*	-0.006	-0.007
	(0.100)	(0.095)	(0.056)	(0.056)	(0.061)	(0.062)	(0.015)	(0.015)	(0.035)	(0.036)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency (ihs)		0.203**		0.031		0.093		-0.008		-0.018
		(0.086)		(0.070)		(0.074)		(0.013)		(0.024)
Pseudo R^2	0.1468	0.1473	0.1489	0.1489	1.211	1.211	1.001	1.001	1.013	1.013
Observations	14,901	14,901	16,956	16,956	16,323	16,323	10,675	10,675	10,675	10,675
Dependent variable mean	0.0847	0.0847	0.0155	0.0155	8,046.3	8,046.3	37,231.1	37,231.1	452.2	452.2
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

Table A12: Average route insecurity and households mobility: robustness to less precisely geocoded events

	Post-harvest departures								**Early season returns**	
	All departures		Temporary		Permanent		Attrition		Return mig. in q1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs)	-0.137	-0.089	-0.161	-0.087	-0.018	0.034	0.214	0.201	-0.120	-0.098
	(0.114)	(0.120)	(0.140)	(0.130)	(0.186)	(0.207)	(0.337)	(0.322)	(0.108)	(0.153)
Orig. conflicts (ihs)	0.157*	0.147*	0.061	0.046	0.333	0.318	1.46***	1.48***	0.115	0.112
	(0.083)	(0.084)	(0.109)	(0.105)	(0.270)	(0.266)	(0.470)	(0.487)	(0.100)	(0.100)
Dest. conflicts (ihs)	-0.026	-0.033	0.128**	0.115*	-0.436***	-0.439***	-0.062	-0.088	-0.037	-0.040
	(0.047)	(0.046)	(0.061)	(0.059)	(0.088)	(0.085)	(0.290)	(0.308)	(0.085)	(0.085)
Orig SPEI drought frequency	0.162*	0.185**	0.084	0.120	0.395***	0.422***	0.120	0.015	0.083	0.091
	(0.084)	(0.085)	(0.088)	(0.102)	(0.088)	(0.075)	(0.225)	(0.233)	(0.088)	(0.091)
Route conflicts (ihs) × Orig SPEI drought frequency		-0.110		-0.198**		-0.096		0.278		-0.041
		(0.071)		(0.085)		(0.157)		(0.394)		(0.117)
Pseudo R^2	0.0664	0.0665	0.0741	0.0743	0.0925	0.0926	0.2340	0.2347	0.1053	0.1054
Observations	180,709	180,709	176,351	176,351	134,758	134,758	4,776	4,776	169,460	169,460
Dependent variable mean	0.0518	0.0518	0.0388	0.0388	0.0186	0.0186	0.0773	0.0773	0.0536	0.0536
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

A.3.5 Robustness to using count of fatalities instead of conflict events

Table A13: Gravity model of internal migration and remittances by drought at origin: robustness with conflict fatalities

	Internal migrants/Stayers		Any transfer		Amount received (FCFA)	
	(1)	(2)	(3)	(4)	(5)	(6)
Route conflict fatalities (ihs), t-1	-0.377***	-0.261**	-0.245***	-0.229***	-0.273***	-0.199**
	(0.141)	(0.125)	(0.066)	(0.066)	(0.097)	(0.094)
Route conflict fatalities (ihs), t-1 × Orig. SPEI drought frequency (ihs)		-0.357***		-0.095		-0.303**
		(0.096)		(0.113)		(0.126)
Pseudo R^2	0.3045	0.3064	0.3268	0.3270	1.061	1.060
Observations	15,294	15,294	13,651	13,651	3,624	3,624
Dependent variable mean	0.0182	0.0182	0.0373	0.0373	4,787.7	4,787.7
Locality-Year fixed effects	✓	✓	✓	✓	✓	✓
Destination-Year fixed effects	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured over a six month period prior to the average regional date of harvest. Origin conflict fatalities includes all fatalities within a 50 km buffer of the locality. Internal migrants/stayers is the average number of migrants reported by the households in the locality for the regions of Mali divided by the number of adult household members. All conflict variables buffers, except the within locality conflict fatalities, exclude a 2km radius around the village. Distance between origin and destination is the length of the shortest route between the two.

Table A14: Average route insecurity and economic migration

	Economic migrants in Mali									
	Internal migrants/Stayers		Any transfer		Amount (FCFA)		Av. dest. wage		Av. dest. dist.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts fatalities (ihs), t-1	-0.098	-0.146	-0.031	-0.035	-0.026	-0.046	-0.002	-0.003	-0.046**	-0.038
	(0.079)	(0.092)	(0.065)	(0.071)	(0.057)	(0.061)	(0.012)	(0.014)	(0.022)	(0.027)
Orig. conflict (ihs), t-1	-0.039	-0.027	-0.024	-0.022	-0.044	-0.038	-0.030**	-0.030**	0.018	0.016
	(0.082)	(0.081)	(0.062)	(0.063)	(0.063)	(0.062)	(0.013)	(0.013)	(0.034)	(0.034)
Orig. SPEI drought frequency, t-1	0.033	-0.117	0.327***	0.314***	0.305***	0.256**	-0.010	-0.011	-0.075***	-0.052
	(0.101)	(0.123)	(0.081)	(0.119)	(0.071)	(0.103)	(0.013)	(0.016)	(0.029)	(0.032)
Dest. conflicts fatalities (ihs), t-1	0.011	0.005	-0.013	-0.015	-0.061	-0.065	-0.028***	-0.028***	-0.004	-0.002
	(0.068)	(0.069)	(0.035)	(0.036)	(0.055)	(0.055)	(0.010)	(0.010)	(0.030)	(0.030)
Route conflicts fatalities (ihs), t-1 × Orig. SPEI drought frequency, t-1		0.121**		0.011		0.042		0.0005		-0.018
		(0.051)		(0.039)		(0.042)		(0.008)		(0.015)
Pseudo R^2	0.1468	0.1473	0.1487	0.1487	1.211	1.211	1.001	1.001	1.013	1.013
Observations	14,901	14,901	16,956	16,956	16,323	16,323	10,675	10,675	10,675	10,675
Dependent variable mean	0.0847	0.0847	0.0155	0.0155	8,046.3	8,046.3	37,231.1	37,231.1	452.2	452.2
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of fatalities during conflict events and measured between the first and second quarter of data collection. Origin conflicts fatalities includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts fatalities, exclude a 2km radius around the village.

Table A15: Average route insecurity and households mobility

	Post-harvest departures								Early season returns	
	All departures		Temporary		Permanent		Attrition		Return mig. in q1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts fatalities (ihs)	-0.071	-0.074	-0.074	-0.042	-0.051	-0.091	0.226	0.211	-0.007	-0.011
	(0.046)	(0.062)	(0.054)	(0.052)	(0.076)	(0.110)	(0.238)	(0.261)	(0.059)	(0.077)
Orig. conflicts fatalities (ihs)	0.075	0.076	0.074	0.070	0.121	0.123	0.779***	0.778***	-0.099	-0.099
	(0.057)	(0.058)	(0.053)	(0.053)	(0.130)	(0.131)	(0.258)	(0.256)	(0.095)	(0.095)
Dest. conflict (ihs)	-0.036	-0.036	-0.013	-0.013	-0.127	-0.133*	-0.216	-0.212	0.018	0.018
	(0.032)	(0.032)	(0.053)	(0.053)	(0.080)	(0.078)	(0.255)	(0.250)	(0.048)	(0.048)
Orig SPEI drought frequency	0.168**	0.165**	0.089	0.121	0.396***	0.347***	0.142	0.111	0.082	0.079
	(0.086)	(0.082)	(0.090)	(0.103)	(0.092)	(0.088)	(0.217)	(0.150)	(0.086)	(0.097)
Route conflicts fatalities (ihs) × Orig SPEI drought frequency		0.005		-0.062*		0.061		0.038		0.006
		(0.037)		(0.037)		(0.071)		(0.225)		(0.050)
Pseudo R^2	0.0664	0.0664	0.0740	0.0742	0.0898	0.0900	0.2358	0.2358	0.1053	0.1053
Observations	180,709	180,709	176,351	176,351	134,758	134,758	4,776	4,776	169,460	169,460
Dependent variable mean	0.0518	0.0518	0.0388	0.0388	0.0186	0.0186	0.0773	0.0773	0.0536	0.0536
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of fatalities during conflict events and measured between the first and second quarter of data collection. Origin conflicts fatalities includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts fatalities, exclude a 2km radius around the village.

A.3.6 Robustness to 25 km buffer for conflicts at origin

Table A16: Average route insecurity and economic migration: robustness to 25km buffer for origin conflicts

	Economic migrants in Mali									
	Internal migrants/Stayers		Any transfer		Amount (FCFA)		Av. dest. wage		Av. dest. dist.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs), t-1	-0.225	-0.333*	-0.081	-0.086	-0.076	-0.105	-0.028	-0.016	-0.067	-0.055
	(0.144)	(0.176)	(0.107)	(0.112)	(0.097)	(0.108)	(0.027)	(0.028)	(0.042)	(0.047)
Orig. conflict 25km (ihs), t-1	0.044	0.065	0.159	0.160	0.208	0.214	-0.069*	-0.071*	-0.034	-0.038
	(0.115)	(0.117)	(0.145)	(0.146)	(0.163)	(0.166)	(0.037)	(0.037)	(0.076)	(0.077)
Orig. SPEI drought frequency, t-1	0.033	-0.112	0.319***	0.309***	0.293***	0.247***	-0.011	-0.0004	-0.078***	-0.063**
	(0.103)	(0.123)	(0.082)	(0.107)	(0.070)	(0.089)	(0.012)	(0.014)	(0.029)	(0.030)
Dest. conflict (ihs), t-1	0.125	0.145	-0.064	-0.064	-0.107	-0.102	-0.037**	-0.039**	0.008	0.007
	(0.113)	(0.109)	(0.063)	(0.063)	(0.077)	(0.078)	(0.017)	(0.018)	(0.043)	(0.045)
Route conflicts (ihs), t-1 × Orig. SPEI drought frequency, t-1		0.271**		0.022		0.102		-0.023*		-0.027
		(0.113)		(0.076)		(0.081)		(0.014)		(0.030)
Pseudo R^2	0.1468	0.1474	0.1489	0.1489	1.211	1.211	1.001	1.001	1.013	1.013
Observations	14,901	14,901	16,956	16,956	16,323	16,323	10,675	10,675	10,675	10,675
Dependent variable mean	0.0847	0.0847	0.0155	0.0155	8,046.3	8,046.3	37,231.1	37,231.1	452.2	452.2
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

Table A17: Average route insecurity and households mobility: robustness to 25km buffer for origin conflicts

	Post-harvest departures								Early season returns	
	All departures		Temporary		Permanent		Attrition		Return mig. in q1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Route conflicts (ihs)	-0.210*	-0.142	-0.208	-0.084	-0.185	-0.151	0.853*	0.822*	-0.095	-0.063
	(0.126)	(0.136)	(0.156)	(0.130)	(0.177)	(0.229)	(0.452)	(0.458)	(0.150)	(0.212)
Orig. conflict 25km (ihs)	0.281	0.272	0.277	0.260	0.358	0.353	1.56**	1.58**	-0.397**	-0.399**
	(0.182)	(0.185)	(0.178)	(0.184)	(0.357)	(0.356)	(0.765)	(0.761)	(0.201)	(0.201)
Dest. conflict (ihs)	-0.086*	-0.090*	0.028	0.017	-0.405***	-0.404***	-0.410	-0.412	-0.058	-0.061
	(0.051)	(0.050)	(0.083)	(0.081)	(0.124)	(0.125)	(0.564)	(0.568)	(0.090)	(0.091)
Orig SPEI drought frequency	0.163*	0.187**	0.084	0.126	0.394***	0.409***	0.187	0.159	0.086	0.095
	(0.084)	(0.085)	(0.089)	(0.103)	(0.088)	(0.074)	(0.227)	(0.190)	(0.086)	(0.089)
Route conflicts (ihs) × Orig SPEI drought frequency		-0.143		-0.298***		-0.060		0.080		-0.054
		(0.095)		(0.103)		(0.193)		(0.464)		(0.139)
Pseudo R^2	0.0666	0.0668	0.0741	0.0745	0.0911	0.0911	0.2271	0.2272	0.1055	0.1055
Observations	180,709	180,709	176,351	176,351	134,758	134,758	4,776	4,776	169,460	169,460
Dependent variable mean	0.0518	0.0518	0.0388	0.0388	0.0186	0.0186	0.0773	0.0773	0.0536	0.0536
Year and Locality fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Estimation with PPML. Standard errors clustered by year and locality are reported in parentheses. *** p>0.01, ** p>0.05, * p>0.1.. The sample period includes all years 2011 and 2014 to 2019. Conflict variables are defined in terms of count of conflict events with at least one fatality and measured between the first and second quarter of data collection. Origin conflicts includes all fatalities within a 50 km buffer of the considered locality. All conflict variables buffers, except the within locality conflicts, exclude a 2km radius around the village.

Chapter II

Cooperation between National Armies: Evidence from the Sahel borders

Joint with Oliver Vanden Eynde

1 Introduction

Borderlands tend to concentrate violence, as evidenced by the fact that 18.7% of recorded violent events in 2023 occurred within 50 km of an international land border, despite these areas accounting for only 5.8% of the world population [1]. Border areas are more prone to violence due to their distinctive political and economic peripheral situations. Additionally, international borders, often coinciding with ethnic boundaries, materialise points of contact between state and non-state actors with potentially diverging interest (Mueller et al., 2022). Beyond these factors, the very discontinuity in state authority at international borders can itself generate conflict dynamics. Security operations are often constrained by international borders. Different countries may not share the same interests in conflicts, and armed groups can use neighbouring territories as a safe haven. Such safe havens have emerged along the border between Afghanistan and Pakistan, or the one between Venezuela and Colombia (Martínez, 2017). However, even if the interests of neighbouring countries are broadly aligned, armed groups could exploit frictions in information-sharing between neighbours, legal constraints on armed forces crossing borders, or a failure to internalise the displacement of conflict across borders. Since 1945, more than 55% of insurgencies that have operated across international borders (Cunningham et al., 2013).

A context that illustrates these frictions clearly is the ongoing Jihadist conflict in the Sahel. This conflict spans several West-African nations and is most intense around the region's porous borders. Particularly, the three-border region, where Mali, Niger, and Burkina Faso converge without distinct physical demarcation, has become the focal point of most violence in the Sahel.

While all major West-African national armies have been involved militarily in a fight against these groups, these unconcerted efforts have failed to contain the proliferation

[1] 31% of recorded violent events in 2023 occurred within 10 km of an international land border, despite these regions accounting for only 23% of the world population

of violence from Mali to neighbouring countries. Our paper investigates if improved cooperation between national armies in the Sahel region makes security operations more effective.

We focus in particular on the creation of a multi-national military force that could cross international borders, known as the 'G5 Sahel' Joint Force. Launched in 2017, this force has prioritized operations in the tri-border region of Mali, Niger, and Burkina Faso. Its primary mandate was to combat terrorism, trans-border crime, and human trafficking [2] Comprising 5,400 personnel from the national armies of Mali, Burkina Faso, Niger, Mauritania, and Chad, the Joint Force conducted regular border patrols and conducted joined and coordinated cross-border operations.

A priori, the effect of this mission on conflict dynamics is unclear. On the one hand, the mission solved legal constraints on operations in border areas, it may have increased troop levels and improved communication between national armies. In addition, it could have led to an internalisation of the externalities that characterize security provision in border areas. On the other hand, the create of the Joint Force might have introduced new coordination frictions between the army units from different countries.

Relying on data from the Armed Conflict Location and Event Database (ACLED), we assess the effect of the introduction of the G5 Joint Force on the basis of two empirical exercises. First, we exploit the limitation of the zone of operation of the G5 Joint Force to a 50 kilometer bandwidth around the concerned borders. Using a regression discontinuity around the zone of operation, we show that conflict is less intense where the G5 mission is active. This result does not seem to be driven by geographical displacement. We observe less violence initiated by Jihadist groups, and less violence initiated by security forces against ethnic militia. As the mandate of the G5 Joint Force was restricted to combating terrorists groups, these results suggest that the G5 mission operated in line with its objectives and achieved a degree of effectiveness in the outcomes we measure.

[2] I Additionally, the Joint Force was tasked with supporting the restoration of state authority and the implementation of development projects and humanitarian operations

Heterogeneity analysis reveals that violence drops more strongly in the operating zone of the G5 Sahel Joint Force where the borders are most porous, i.e. when the same ethnic group is present on both sides of the border and the border does not align with a major river.

To shed more light on the underlying mechanism, we conduct a second exercise in which we try to assess more directly if the G5 mission facilitated security operations in border areas. We focus on major French attacks on jihadist groups. We expect these attacks to trigger the movement of jihadist groups and new security operations. In the aftermath of these trigger attacks, we do see less security operations in border areas when the G5 mission is not active. However, when the G5 mission is active, this effect disappears. These findings offer additional support for the idea that the G5 force facilitated security operations in border areas.

Our paper adds to a small literature in economics and political science that studies the effect of borders on conflict outcomes with granular conflict data. Martínez (2017) shows that the presidency of Hugo Chavez in Venezuela increased the presence of FARC rebels in Colombian municipalities along the border. Studying the geography of conflict, Mueller et al. (2022) argue that raising physical barriers at ethnic frontiers could reduce conflict. Blair (2023) provides evidence from Iraq showing how border protections reduce the victimization of civilians by rebel fighters. However, this paper does not look at how borders affect the propagation of groups whose objectives are not limited to one country. In addition, the role of a cross-national force is particularly interesting as a policy intervention, as the construction of fences is practically not feasible in many settings, and may hamper with economic activity in border areas.

To our knowledge, our paper is the first to measure the effect of military cooperation between national armies on the conflict dynamics in border areas. We add to the large literature on the empirical study of conflict using econometric methods. This literature has studied the effects of economic shocks on conflict extensively (e.g. Miguel et al., 2004;

Ferrara and Harari, 2018; Dube and Vargas, 2013; Berman et al., 2015; Vanden Eynde, 2016). The role of religious and ethnic diversity (e.g. Montalvo and Reynal-Querol, 2005; Esteban et al., 2012), as well as the role of political institutions as drivers of civil conflict (Besley and Persson, 2011) have also been studied at length.[3] In parallel, there is increasing evidence on how development interventions affect conflict (Berman et al., 2011; Crost and Johnston, 2014; Fetzer, 2020)). The role of media and information interventions are also increasingly studied. For example, (Armand et al., 2020) find that radio campaigns can contribute to demobilization of armed groups.[4] Finally, and closest to the current project, a very recent literature evaluates the effect of military interventions. For example, Dell and Querubin (2018) find that aerial bombing campaigns by the US in Vietnam increased the support for communist insurgents. There is not much work studying in the organizational aspects of war and military planning.[5] Fetzer et al. (2021) study changes in military cooperation in the context of the security transition from NATO to the Afghan National Army in Afghanistan. This paper finds that the security transition improved security in a first stage, but worsened outcomes when NATO troops were withdrawn physically. The authors argue that these patterns are consistent with the Taliban lying low strategically to facilitate of the withdrawal. A related study shows that frictions between different NATO allies led to worse security outcomes Fetzer et al. (2024). Our paper adds policy relevance to this work, by showing that institutionalized cooperation between armies can impact security outcomes. There is a broader question under which conditions external military interventions can be effective (For a recent survey,

[3]A number of recent papers have highlighted how specific sub-national institutions can spur or mitigate conflict (e.g. Shapiro and Vanden Eynde, 2023; Fetzer and Kyburz, 2023), or how institutional arrangements arise in war settings (Sanchez de la Sierra, 2020; Dincecco et al., 2022).

[4]For related work on the role of media, see: Yanagizawa-Drott (2014); Durante and Zhuravskaya (2018); Adena et al. (2015). While communication technology touches on a very important aspect of war, the focus of this existing work is on persuasion or coordination outside of the security forces. In contrast, this paper focuses on military cooperation.

[5]Exceptions are Ager et al. (2022), who study the role of incentives for fighter pilots in the German air force during World War II, and Acemoglu et al. (2020) who study the incentive for Colombian soldiers to target civilians and claim them as rebel fatalities. However, the focus of our paper is not on individual incentives, but on military cooperation.

see Rohner, 2024). In the African continent, there is evidence military UN peacekeeping missions help to reduce conflict (Hultman et al., 2014). They also appear to protect the civilian population against rebel abuse, but not against abuse by government forces (Fjelde et al., 2019). Of course, these settings are distinct from the one of military cooperation between neighbouring countries which we focus on in the current paper. Given the importance of international alliances for the effective provision of security, the questions addressed by our paper are particularly relevant - for the Sahel region and beyond.

As far as work on the Sahel region is concerned, our paper is also one of the first quantitative empirical studies of conflict in the region. Focusing on the seasonal migration of herders (transumant pastoralists), McGuirk and Nunn (2022) find that rainfall deficiency has exacerbated the conflict between pastoralists and agriculturalists. In line with these findings, Eberle et al. (2020) show that temperature increases generate more violence between farmers and herders. While these two recent papers do not explicitly restrict their analysis the Sahel region, the processes they study are particularly important in the region we study. In Mali, Richard (2022) finds that the insecurity induced by the conflict hampers seasonal migration and, hence, reduces lean season consumption in village usually relying on this type of migration and an income source. Calvo et al. (2020) study the effect of conflict in Mali on social capital. They find that conflict exposure increases engagement political associations, which could deepen the conflict to the extent that these organizations act as interest groups for particular ethnic groups. Premand and Rohner (2023) study a large-scale conditional cash transfer scheme in Niger, and find that this programme increased conflict intensity. These recent contributions all shed light on important aspects of the conflict. However, our paper is the first to focus on the security operations in this conflict.

2 Background

The Sahel region has been plagued by conflict in recent years, with armed groups operating across borders and increasing violence since the 2012 Tuareg-led rebellion in Mali. This rebellion was followed by the proliferation of armed groups, including ethnic militias and jihadist groups. The concentration of violent events has been particularly high in the three-borders area, spreading from Mali to neighboring Niger and Burkina Faso.This region is inhabited by various ethnic groups, including the Fulani and Tuareg, who constitute a majority in the area but are ethnic minorities in their respective home countries. As a result, it has become a recruitment target for jihadist groups. Over the past year, 31% of conflict-related fatalities in these countries occurred within 50 kilometers of these unmarked borders [6].

To address this instability, several international peacekeeping and counter-terrorism missions have been launched in the region. In January 2013, the French launched Operation Serval to regain control of the north of Mali, which was followed by the Barkhane operation, which has maintained a lasting presence in the country until its complete withdraw in May 2022. In April 2013, the United Nations deployed the Multidimensional Integrated Stabilization Mission in Mali (MINUSMA).

In February 2014, the G5-Sahel was created to facilitate military cooperation between Mali, Burkina Faso, Niger, Mauritania, and Chad. The G5-Sahel Joint Force was established by a UN resolution in June 2017,[7] and its command became active in September of that year. While the G5S initiative was strongly supported by France, analysts underline that the initiative came largely from the participating countries themselves (Touchard, 2018). Morevoer, the G5S force was operationally independent from the French military mission, altough some operations were conducted in coordination with Barkhane [8], and reports from meetings of the highest decision instance of the organisation report partici-

[6] 56.7% within 100 kilometers, 10.1% within 10 kilometers
[7] UN Resolution 2364 (2017).
[8] ACLED

pation of Chief of Staff of the French Armed Forces and the commander of the Barkhane force operating in the Sahel [9]. Furthermore, the UN mission in Mali played an essential role for the local logistical support of the G5 Joint Force.

The G5-Sahel force had a mandate focused on combating terrorist groups and trafficking gangs in border areas. It gave national armies the possibility to operate outside of the national territory, within pre-defined operation zones. In those zones, the force was also tasked with restoring state authority, and supporting development and humanitarian interventions (Touchard, 2018).

The Joint Force of the G5 Sahel was composed of 5,400 men across eight battalions, comparable to the size of the Barkhane force. The battalions were distributed around three operations areas, with a Malian and a Mauritanian battalion operating jointly in the Western operation zone at the border between Mali and Mauritania, four battalions with troops from Mali, Burkina Faso, Niger and Chad based in the Central Zone, at the border between Mali, Niger and Burkina Faso (the three-borders area), and a Nigerian and Chadian battalion based in the Eastern operation zone at the border between Chad and Niger. In G5 Sahel internal unclassified documents we have access to, 14 operations were reported in the Central Zone between 2017 and 2020, while only five and four operations were conducted respectively in the Western and Eastern operation zone during the same period.

The planning of operations, along with other operational responsibilities such as coordinating resources from member countries and ensuring effective communication and cooperation between operational zones, fell under the purview of the Commander of the Joint Force, and had to be approved by the Council of Defense and Security (CDS), composed of the Chief of Staff of national Armed Forces of its members, and meeting on a regular basis twice per year. However, internal reports from the G5 Sahel suggest that, in practice, heads of national armed forces retained control over operational

[9] G5 Sahel, 24 October 2019 https://cdg5s.org/fr/node/1335

responsibilities and often engaged in bilateral dealings rather than utilizing G5 Sahel institutions. Despite these shortcomings, the G5S mission facilitated communication between army units along the border. For example, reports mention the creation of a phone book covering the border zones with mobile numbers of officers along the border, which improves the reactivity of forces in these areas (Boeke and Chauzal, 2017).

For each of the three zones of operation, the G5S had a dedicated command. Commanders of these operation zones were in charge of tactical decisions, such as planning routine operations such as patrols or conducting spontaneous operations either to exploit a temporary vulnerability of an terrorist groups (called "opportunistic operations"), or respond to urgent needs of civilian protections.

Among the operations conducted by the Joint Force listed in in its internal documents or documented in the ACLED data were regular patrols in border areas, as well as spontaneous or planned joint cross-border operations. These operations involved airstrikes on armed group positions, conducted in cooperation with Barkhane, as well as the neutralization or arrest of terrorists and seizure of weapons.

The operating zones for the G5-Sahel mission were initially defined as buffers of 50 km around the international borders the borders, but in January 2020, an announcement was made to extend the operating zone to 100 km.[10] As the implementation of this extension is unclear,[11] the focus in our empirical analysis remains on the 50km buffer zone in the time period between September 2017 and January 2020.

We do not have access to the exact geographical locations for all G5-Sahel operations operations. However, in figure A 1 we map all the events reported in the ACLED data that mention the G5-Sahel force explicitly or that involve G5-Sahel members operating outside of their national territories. The spatial distribution of these events is consistent

[10] G5 Sahel, 26 January 2020, https://www.g5sahel.org/les-chefs-d-etat-major-des-pays-du-g5-sahel-rendent-plus-operationnelle-la-force-conjointe.

[11] We observe not a single trans-border operation outside of the 50km buffer in the ACLED data. As late as December 2021, the French UN delegation website described the G5 mission as being active in a 50km zone (https://onu.delegfrance.org/france-s-action-in-the-sahel.)

with a focus of the G5 Sahel joint force on the three-borders region between Mali, Niger and Burkina Faso, and with a restriction of operations within the 50 km buffer zone around these borders. For most of our empirical analysis, we will hence focus on this Central zone.

The activities of the G5 Sahel force were temporarily disrupted in June 2018, when a major suicide attack hit the G5 Sahel Joint Force headquarters. The resulting damage led to a suspension of activities until January 2019. The most active period of the G5-Sahel seem to have been between Agust 2019 and July 2021 under the command of General Oumarou Namata Gazama, Deputy Chief of Niger's Army, with 11 large-scale operations conducted during the period, besides routine and opportunistic operations [12]. From 2020 onward, two military coups in Mali generated tensions with the international community. The country finally left the G5-Sahel in May 2022, leading to the de facto end of the organization. Nevertheless, the G5-Sahel Joint Force was an essential part of international efforts to restore stability to the region for more than 5 years. Our paper attempts to assess its effect on conflict dynamics.

3 Theoretical Framework

The provision of security in border areas entails specific challenges. Armed groups may be more mobile than security forces, who are not allowed to cross borders. Anticipating the possibility of armed groups to flee to safe havens, security forces may decide not to intervene in border zones.[13] In addition, investments in security in border areas may have externalities for neighbouring countries which are not taken into account when countries decide on how much investments to make in border zones. For all these

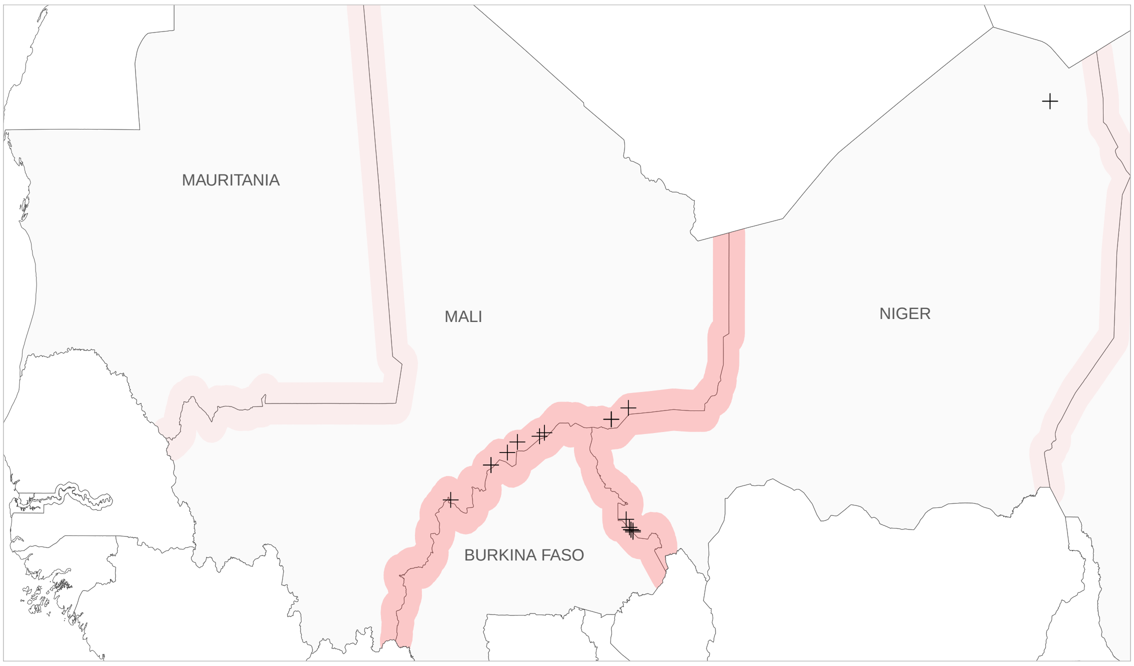
[12]Source: Closing statement of the term of General Oumarou Namata Gazama as Commander-in-Chief of the G5 Sahel Joint Force, 31st July 2021

[13]Theoretically, the possibility of displacement could also lead to a number of security operations in border zones that is above its optimal level, as neighbouring countries compete to try to push groups across the border. However, this mechanism seems most plausible when investments in security and displacement are long-lived, which is not the case in the context of the Sahel.

Figure 1: Foreign Military Operations in Sahel regions

Panel A: Foreign Military Operations before G5-Sahel Joint Force Creation

Panel B: Foreign Military Operations during G5-Sahel period



Notes: Foreign Military Operations include all events involving G5 Sahel or G5-Sahel country members military forces outside of their national territory. The red shaded areas represent the 50 km operation zones around borders where the G5-Sahel operated officially between January 2017 and January 2020.

reasons, security provision might be under-optimal in border areas.

The G5-Sahel mission could have increased or decreased conflict intensity, depending on the mechanism of influence. First of all, counter-terrorism forces could be a deterrent or neutralize armed groups, leading to a reduction in observed violence. However, counter-terrorism operations are also likely to result in greater fatalities among armed group members, and they could lead to violent escalation. Similarly, counter-terrorism forces could be an additional risk factor for civilians. These may become collateral victims, or could suffer from retaliation of armed groups against suspected collaborators. Moreover, organizational frictions between national armies could reduce the effectiveness of the joint force. Descriptive accounts point at severe coordination problems in terms of equipment and command structures (Touchard, 2018).

In thinking about the impacts of the G5 mission, the effects could be direct, involving the operations of G5-Sahel units, but also indirect when they involve other units whose behaviour changes in response to the presence of the G5-Sahel mission. For example, the improved communication between border forces that the G5S mission facilitated is likely to affect all units.

Our empirical approach will not enable a fine distinction between all mechanisms. However, the sign of the net effects we estimate, in combination with a detailed analysis of different types of violence, will help us to narrow down the mechanisms underlying our findings. In addition, heterogeneous effects can shed some further light on the causal channels. If the G5 mission improves military coordination in border areas where militants are more mobile, we could expect the impacts to be most pronounced where it is also easier for armed groups to move without being detected or interrupted – for example in rugged areas.

4 Data

Our main source of violence data is ACLED (Armed Conflict Location Event Data). ACLED is a database that tracks and records information on armed conflicts and political violence around the world. The data is sourced from a variety of sources, including traditional media, social networks, NGOs, international organizations, and local partners. For each violent event, the ACLED data records the number of fatalities that occurred. The events are also geo-coded, meaning that their location is identified and mapped. The locality of the events is coded at different levels of precision. This coding precision will be important for our study, as our regression discontinuity approach will exploit fine geographical variation. In addition to its localisation, each ACLED event is dated precisely. Further sub-categories are created based on the actors involved in each event. These sub-categories help to differentiate between different types of violence, such as conflicts between state and non-state actors or violence between different non-state actors. Overall, the ACLED data provides a comprehensive and detailed record of violent events and their associated actors, fatalities, and locations. It is important to acknowledge that such data may be subject to biases and limitations. While ACLED relies on various sources to compile its database of conflict events and casualties, this reliance on often fragmented and incomplete information can lead to underreporting or overreporting of casualties, particularly in areas with limited media coverage or where access is restricted. While the reported fatalities are likely to be biased as underscored by the ACLED data description as ”prone to manipulation by armed groups, and occasionally the media”, this noisy measure may still contain some information about conflict intensity, and has been shown to correlate strongly with climatic shocks (e.g. Ferrara and Harari, 2018), population displacement (e.g. Tai et al., 2022). and child health (e.g. Tapsoba, 2023). As a robustness check, we also use a transformation of the fatality variables recording whether any event with any fatality is reported.

We map the ACLED data and use information on the operation zones of the G5-

Sahel mission that we obtained from official documents. To this mapped violence data, we add granular data on nightlight emissions and geographical features such as road access, urbanization, ruggedness and closeness to rivers. These additional variables will be used to support the validity of our empirical approach.

Figure A 1 shows the trends in violence, based on the ACLED data, in the ”central operation zone”, i.e. the three border area of Burkina Faso, Mali, and Niger. Violence is clearly trending upwards from early 2017 onwards. This increase is particularly pronounced in areas very close (within 15 km) to the international borders, while the broader operation zone of the G5 mission follows a trend that is similar to the one observed for the areas outside of the G5 operation zone. While this graph illustrates the overall conflict dynamics and is important for the context of our study, we do not think it allows us to identify the effect of the G5-Sahel mission. The low levels of violence before the launch of the G5-Sahel force in September 2017 make the setting ill-suited for a difference-in-difference approach. Moreover, the G5-mission was created in anticipation of the conflict becoming more gradually intense in the border areas, and the strategies of various actors may have contributed to this intensification. Hence, estimating the effect of cooperation between national armies necessitates a more granular approach. We develop such empirical strategies in the next section.

To characterize borders' porosity as well as potential logistical support available to the Joint Force, we use three additional data sources on the spatial distribution of ethnic groups from Desmet et al. (2020), on major rivers flows using the river network in Africa produced by the World Agro-forestry center, and on the location of UN-Peacekeeping Missions in Mali using The Geo-PKO dataset (Cil et al., 2020).

5 Empirical strategy

We will study the effect of the G5 Sahel mission through the lens of two empirical exercises.

5.1 Discontinuity around G5 operation zones.

First, we will use a regression discontinuity design to assess whether violence levels are different in the operation zones of the G5-mission. We are interested in comparing areas where national armies cooperate to areas that are under the sole responsibility of national army units. The precise definition of the operation zone of the G5 mission offers a plausibly exogenous assignment to these two security environments. In particular, we leverage the spatial discontinuity created by the limitation of the G5-Sahel operation zone within 50 km of G5-Sahel countries' borders:

$$\begin{aligned} y_i = \alpha + \beta Border_i + \delta\left(Distance_i - z_0\right) \\ + \delta' Border_i \cdot \left(Distance_i - z_0\right) + \delta_b + \eta_c + \epsilon_i \end{aligned} \tag{1}$$

In this specification, the outcome y_i is our measure of conflict at the grid-cell i level. z_0 refers to the limit of the buffer zone, at 50 km. $Distance_i$ measures the distance to the 50km buffer limit. We could expect this running variable to correlate with conflict outcomes - for example, conflict could be systematically more intense when we are closer to the international border. In our empirical approach, we want to control for such impacts, and evaluate instead whether the operation zone of the G5 creates a discontinuous change in conflict outcomes. The discontinuity is be captured in the equation by $Border_i$, which indicates whether grid-cell is less than 50km from the border. We also include border segment fixed effects δ_b and country fixed effects η_c. We allow for a data-driven choice of two bandwiths for optimal mean squared error (MSE) point estimation with

Calonico-Cattaneo-Titiunik procedure (Calonico et al., 2014). As the RD approach relies on a fine coding of conflict events, we use granular gridcells (0.025 by 0.025 degrees), and we focus on ACLED events with the highest precision level for our main results. We show findings for alternative coding as robustness checks. To support the validity of the RD approach, we will show continuity of geographical characteristics and pre-G5 levels of violence around the border of the operation zone.

5.2 Response to trigger events

To shed further light on the mechanisms underlying the RD results, we will rely on a second empirical exercise. Here, we study how the response of violence to trigger events differs in border areas and depending on whether the G5 mission is active or not. As trigger events, we focus on major French operations against Jihadist groups, which we identify as operations that claim at least 5 fatalities. These events are followed by a marked intensification of conflict, as shown in figure A 6 in the online appendix. In a triple difference approach, we will compare the response to trigger events between border and non-border areas, and when the G5-Sahel mission is active versus not active. As for the RD results, we use 0.025 by 0.025 degree grid-cells. We restrict ourselves to trigger operations that are less than 250km from the three-border area. Then, we construct a window of 8 weeks,[14] and a geographical circle of a 100km radius, around each trigger operation. Figure 2 illustrates the spatial definition of trigger operations and border areas.When gridcells are part of multiple event windows, we keep only the first window for that gridcell.[15] In total, our sample has 48 trigger operations, and 20 of these take place when the G5-Sahel mission is not active.

[14]We exclude the day of the trigger event and the day after, to avoid mechanical effects.

[15]This approach prevents us from using already treated gridcells as a control in future comparisons. As has been highlighted by the recent literature on difference-in-difference methods, such comparisons could introduce negative weights in the estimated treatment effect (de Chaisemartin and D'Haultfoeuille, 2022).

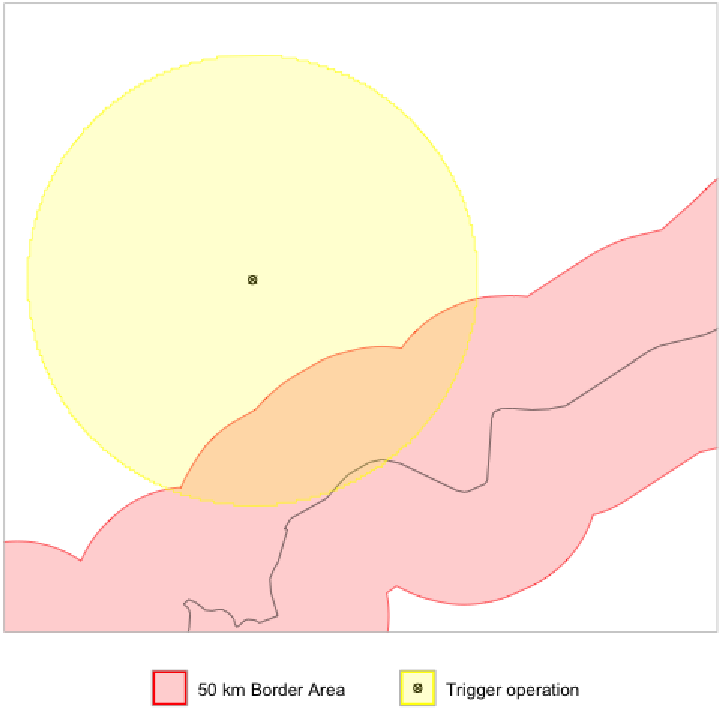


Figure 2: Definition of trigger response areas and border areas

The resulting estimating equation is:

$$y_{i,p,t} = \zeta Border_i * Post_{t,p} + \chi Border_i * Post_{t,p} * G5_t + \eta_i + \gamma_{p,t} + \epsilon_{i,p,t} \tag{2}$$

In this equation, $y_{i,p,t}$ is a measure of conflict at time t in gridcell i for a window around operation p. t is measured in two-week periods. The outcome $Border_i$ indicates whether grid-cell is less than 50km from the international border, so within the operation zone of the G5-Sahel force. $Post_{t,p}$ is one in the time periods after the trigger operation. $G5_t$ is an indicator for when the G5 Force is active.[16] We also include operation by time fixed effects ($\gamma_{p,t}$), as well as gridcell level fixed effects (η_i). We cluster standard errors at the gridcell-level. While our main interest is in the triple difference specification above, we

[16] As we are interested in immediate response patterns, we will consider the period in which the G5-Sahel mission was incapacitated by a suicide attack on its headquarters as a period in which the G5 is not active, between June 2018 and January 2019.

will also show results where we estimate time-to-operation effects in event study graphs.

For both empirical exercises, we consider the period of activity of the G5 Sahel between September 2017 and January 2020 within 50 km of border areas, at the zone of operation of the Joint Mission after January 2020 and the exact date of the end if its activities cannot be defined clearly. For the response analysis, as we are interested in immediate military response to initial trigger events, we consider the period between June 2018 and January 2019 in which the G5-Sahel mission was incapacitated by a suicide attack on its headquarters as a period in which the G5 is not active, but check that the results are robust to considering this period as active.

6 Results

6.1 RD results

Figure 3 compares fatalities in gridcells within the G5-Sahel central operation zone to those in gridcells just outside the operation zone during the period of activity of the G5 mission within 50km of international borders. The local linear regressions in Panel A and B shows a clear discontinuity in the number of total conflict fatalities and civilians fatalities caused by Islamist groups at the border of the operation. There are less fatalities where the G5 mission is active, especially for events involving Islamist groups whose containment was the objective of the G5 mission. The discontinuity appears stronger with the data-driven optimal bandwidth (Calonico-Cattaneo-Titiunik procedure to minimise bias and variance of the RD estimator) which mostly excludes cells closer to the international border.

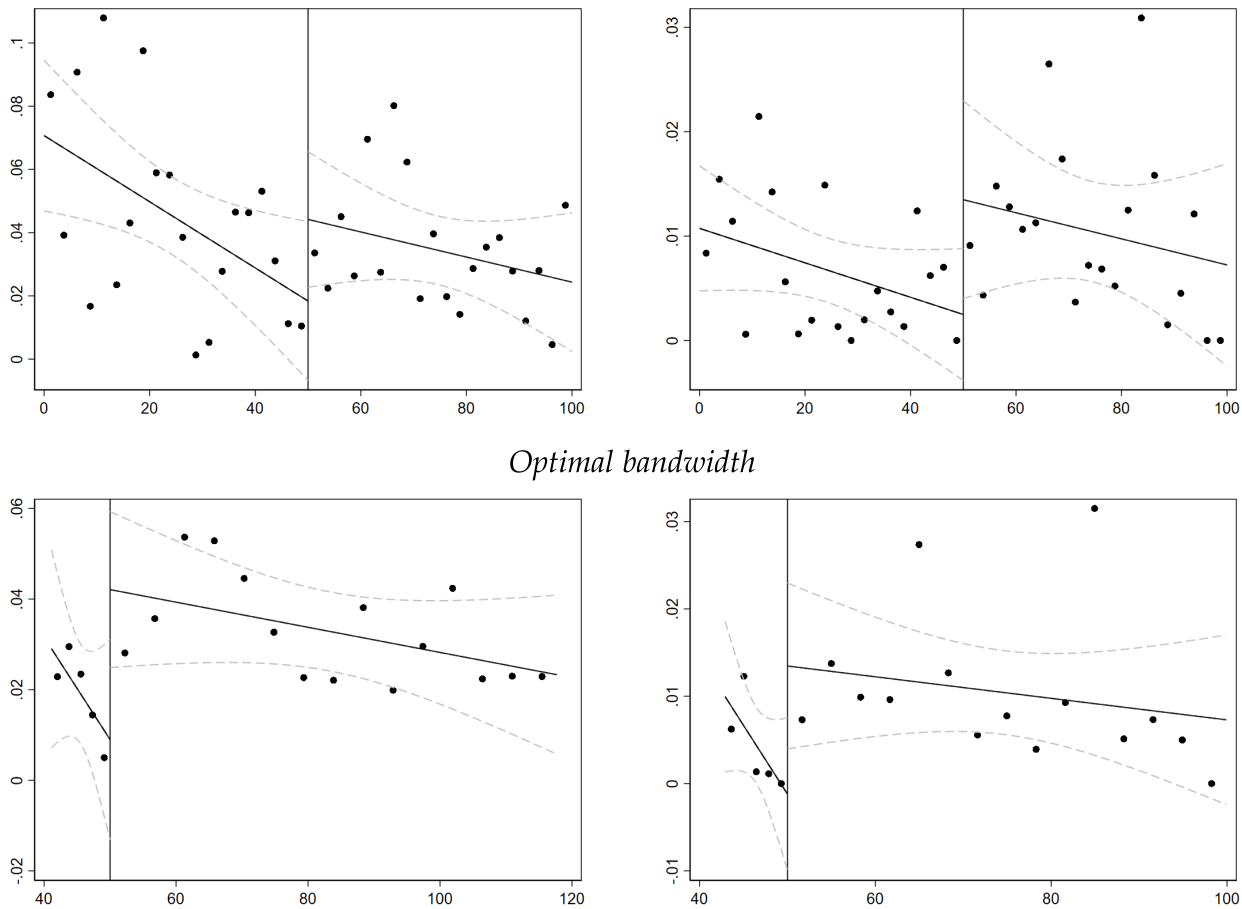
The graphical patterns show a gradual increase in conflict as one moves closer to the international border. This effect can be due to the strategic importance of the border for the security forces, as well as the net benefits of operating in these areas from the perspective of armed groups. However, the RD results show a discontinuous jump in this pattern within the operation zone of the G5. The graphical patterns also suggest that the results are not driven by symmetrical displacement, where the reduction within the operation zone is offset by increased violence just outside the operation zone. Of course, it is impossible to rule out more gradual displacement. However, interpretationally, this would still provide evidence of the local effectiveness of the G5-Sahel mission. As we set out in the theoretical framework, it is far from obvious ex ante that the mission should have reduced violence - even locally.

Figure A 2 in Appendix shows the comparison for the total number of fatalities during the G5 mission period and before September 2017, when the G5 mission was not active. There is no discontinuity in the pre-treatment period, which supports the validity of the RD approach.

Figure 3: Regression Discontinuity for G5 central operation zone

Panel A: All events fatalities *Panel B:* Attacks on civilians by Islamist groups

Symmetric full bandwidth



Notes: Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 200km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. The sample comprises the central Zone (Mali-Niger-Burkina Faso three borders regions). Included events are coded with geo-precision level 1.

Table 1: Discontinuity in conflict intensity for G5 Sahel Central operation zone by actors 2017-2020.

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All events								
Robust	-0.0449***	-0.0108	-0.0050**	-0.0022***	-0.0359**	-0.0073	-0.0319***	-0.0033
	(0.0165)	(0.0073)	(0.0025)	(0.0007)	(0.0151)	(0.0069)	(0.0113)	(0.0051)
Mean DV	0.027	0.011	0.003	0.001	0.024	0.010	0.016	0.007
Standard Deviation	1.115	0.220	0.211	0.033	1.073	0.198	0.646	0.123
Observations within buffer	3726	5881	7285	4865	3609	6378	3285	7668
Observations untreated	38392	25505	27652	42229	32315	22713	34083	22389
Bandwidth untreated (km)	71.101	46.514	50.720	78.387	59.485	41.264	62.884	40.734
Bandwidth treated (km)	6.499	10.214	12.582	8.413	6.280	11.000	5.684	13.220
Panel B: Events involving Islamist groups								
Robust	-0.0244**	-0.0094***	-0.0015	-0.0008	-0.0244**	-0.0094***	-0.0132***	-0.0041*
	(0.0115)	(0.0035)	(0.0017)	(0.0005)	(0.0115)	(0.0035)	(0.0047)	(0.0021)
Mean DV	0.015	0.006	0.000	0.000	0.015	0.006	0.005	0.003
Standard Deviation	0.875	0.123	0.067	0.011	0.875	0.123	0.273	0.076
Observations within buffer	6957	5243	8599	6298	6957	5243	4079	8771
Observations untreated	37374	39015	30246	28692	37374	39015	25485	45587
Bandwidth untreated (km)	69.241	72.389	55.643	52.692	69.241	72.389	46.456	84.874
Bandwidth treated (km)	12.079	9.105	14.864	10.868	12.079	9.105	7.090	15.092
Panel C: Events involving Communal militia								
Robust	-0.0084	0.0019	-0.0025	-0.0017***	-0.0084	0.0019	-0.0081	0.0035
	(0.0085)	(0.0033)	(0.0018)	(0.0006)	(0.0085)	(0.0033)	(0.0076)	(0.0029)
Mean DV	0.008	0.002	0.002	0.001	0.008	0.002	0.006	0.001
Standard Deviation	0.573	0.058	0.195	0.028	0.573	0.058	0.516	0.042
Observations within buffer	3270	11604	6942	4750	3270	11604	3114	12034
Observations untreated	39963	24751	29184	34175	39963	24751	35309	20710
Bandwidth untreated (km)	74.163	45.224	53.602	63.075	74.163	45.224	65.131	37.586
Bandwidth treated (km)	5.664	19.825	12.034	8.208	5.664	19.825	5.421	20.548
Panel D: Events involving civilians								
Robust	-0.0319***	-0.0033	-0.0020*	-0.0011***	-0.0249**	-0.0004	-0.0319***	-0.0033
	(0.0113)	(0.0051)	(0.0012)	(0.0004)	(0.0104)	(0.0046)	(0.0113)	(0.0051)
Mean DV	0.016	0.007	0.002	0.000	0.013	0.006	0.016	0.007
Standard Deviation	0.646	0.123	0.191	0.024	0.600	0.112	0.646	0.123
Observations within buffer	3285	7668	7553	5101	3190	9994	3285	7668
Observations untreated	34083	22389	16038	33412	27594	21360	34083	22389
Bandwidth untreated (km)	62.884	40.734	28.872	61.633	50.602	38.746	62.884	40.734
Bandwidth treated (km)	5.684	13.220	13.036	8.807	5.544	17.149	5.684	13.220

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of borders between Burkina Faso, Mali and Niger are included for the period between January 2011 and September 2017. Include events coded with geo-precision level 1 only. Empty cells means the discontinuity could not be estimated due to lack of variability in the dependent variable. Military operations involving islamist groups record fatalities from islamist groups only, whereas attacks by armed group and attacks on civilians involving islamist goups record fatalities caused by islamist groups. Idem for communal militia and unidentified armed groups. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2: Continuity in conflict intensity and geographical variables for G5 Sahel Central operation zone before G5 Sahel first operation.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Conflict variables before Sept 2017		Geographical variables							
Variables:	Events	Fatalities(#)	GHS pop	Nightlight 2017	Nightlight 2014	NDVI	Road	Rivers	Cities	Ruggedness
Conventional	0.0005	-0.0004	-0.3384	-0.0034	-0.0024	0.2634	-0.3823	0.2758	0.1124	0.0909
	(0.0021)	(0.0020)	(4.1348)	(0.0027)	(0.0029)	(0.2451)	(0.4188)	(0.6033)	(0.6682)	(0.4954)
Bias-corrected	0.0007	-0.0006	-2.2285	-0.0044*	-0.0032	0.2521	-0.6176	0.2674	-0.0362	0.0989
	(0.0021)	(0.0020)	(4.1348)	(0.0027)	(0.0029)	(0.2451)	(0.4188)	(0.6033)	(0.6682)	(0.4954)
Robust	0.0007	-0.0006	-2.2285	-0.0044	-0.0032	0.2521	-0.6176	0.2674	-0.0362	0.0989
	(0.0024)	(0.0028)	(4.7529)	(0.0032)	(0.0035)	(0.2928)	(0.4756)	(0.7200)	(0.7954)	(0.5605)
Mean DV	0.004	0.010	36.676	0.195	0.009	131.669	21.677	36.819	116.643	13.787
Standard Deviation	0.357	0.920	214.323	0.262	0.260	16.791	36.265	33.116	101.818	20.645
Observations within buffer	5369	4476	7415	10155	6911	6319	8565	11255	12201	6700
Observations untreated	14332	6874	14654	11793	11971	15691	9974	17162	18665	8946
Bandwidth untreated (km)	25.662	12.131	26.330	21.065	21.399	28.188	17.742	30.911	33.761	15.923
Bandwidth treated (km)	9.389	7.804	12.804	17.466	11.981	10.904	14.798	19.211	20.821	11.560

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of borders between Burkina Faso, Mali and Niger are included for the period between January 2011 and September 2017. Include events coded with geo-precision level 1 only. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Panel A in Table 1 shows RD estimates for different estimation approaches and outcomes. We see that the general pattern of lower fatalities in the G5 operation zone holds across estimation methods and outcomes. The magnitude of the effect is large: crossing into the operation zone of the G5 mission reduces the number of fatalities in a given gridcell by 0.04, whereas the mean number of fatalities per gridcell is around 0.03. The reduction in violence is observed for security operations, attacks by armed groups, and violence against civilians. The effects are a bit more marked for the number fatalities than for event counts, but they go in the same direction.

To shed more light on the mechanisms underlying the observed reduction in violence, the subsequent panels in Table 1 show RD results for a finer classification of violent events based on the actors involved. Panel B focuses on violence involving islamist groups. Interestingly, security operation against islamist groups do not decrease significantly in the operation zone of the G5 region. However, there are less attacks by armed jihadist groups and less attacks against civilians. Strikingly, there is a more pronounced reduction in security force violence against ethnic militia groups, while these ethnic militias do not reduce their violence significantly. As the official mandate of the G5-Sahel force is focused on combating Jihadist groups, these findings suggest the G5-Sahel mission reduces violence initiated by the actor it is supposed to target. In this sense, the mission appears to be effective.

Table 2 presents important validity checks for the RD approach. It confirms the absence of discontinuities in pre-G5 conflict measures as well as geographical characteristics. Hence, we are confident that the regression discontinuity estimates are picking up the causal effect of grid-cells belonging to the operation zone of the G5 mission.

Heterogeneity analysis

Table 3 presents heterogeneous results based on the characteristics of the border. The observed reduction in violence is particularly prominent in border segments aligned with rivers and where a shared ethnic group resides on both sides of the border. Rivers serve as natural barriers, enhancing detection and interception capabilities, while the presence of a common ethnic group facilitates cross-border mobility and operations of armed groups. Consequently, these findings support the hypothesis that the effectiveness of the G5 is greatest in regions with porous borders. [17] Figure A 5 in appendix illustrates the definition of border segments characteristics used for this heterogeneity analysis. The patterns are a bit harder to interpret for the distance to MINUSMA forces. The reduction in violence is larger in the vicinity of MINUSMA forces, which could capture complementarities between the G5 Sahel mission and MINUSMA.

Table A1 in appendix further shows heterogeneity by border and country member, suggesting the G5 is most effective at the border between Mali and Burkina Faso, and to a lesser extent at the border between Mali and Niger, but does not seem to affect significantly conflict intensity at the border between Niger and Burkina Faso. At country level, the strongest effect is detected for Burkina Faso, and then for Mali. These results together suggest that the G5 may be most effect at containing conflict whithin Burkina Faso along the border with Mali. This in line with the characteristics of these borders describe in Table A6: the border between Mali and its two neighbors tend to be closer to MINUSMA bases, and have greater presence of transborder ethnic groups. These characteristics may, in turn, may drive a greater "Mali effect". We cannot test formally however, to which extent these factors drive the observed heterogenous effects across borders. The second part of Table A6 in turn, show that among G5 operations observed in the ACLED data (represented in Figure 1), there were relatively more operations conducted within Mali, and especially at the border between Mali and Burkina-Faso. While reporting issues may affect the reliability of the ACLED data on foreign interventions, these patterns are consistent with a general objective of the G5 to prevent the spread of insecurity from Mali to its neighbors.

[17] Modern RD methods are not set up to estimate heterogeneous effects, so we do not formally test the difference between coefficients. Another caveat to this interpretation is that that some of difference between these regions already appears in the mean of the dependent variable.

Table 3: Discontinuity in conflict intensity for central G5-Sahel operation zones 2017-2020, heterogeneity by border segment characteristics.

	Low rugged		High rugged		River		No river		Close MINUSMA		Far MINUSMA		No common group		Transborder group	
	Fatalities (1)	Events (2)	Fatalities (3)	Events (4)	Fatalities (5)	Events (6)	Fatalities (7)	Events (8)	Fatalities (9)	Events (10)	Fatalities (11)	Events (13)	Fatalities (11)	Events (14)	Fatalities (15)	Events (16)
Conventional	-0.0306	0.0064	-0.0489***	-0.0242***	-0.0561	0.0016	-0.0409***	-0.0166**	-0.0956**	-0.0096	-0.0089*	-0.0100**	-0.0332**	-0.0059	-0.0647**	-0.0193**
	(0.0194)	(0.0089)	(0.0186)	(0.0072)	(0.0484)	(0.0151)	(0.0119)	(0.0066)	(0.0402)	(0.0143)	(0.0048)	(0.0041)	(0.0155)	(0.0087)	(0.0287)	(0.0084)
Bias-corrected	-0.0403**	0.0079	-0.0442**	-0.0256***	-0.0687	0.0022	-0.0411***	-0.0182***	-0.1084***	-0.0105	-0.0085*	-0.0107***	-0.0382**	-0.0076	-0.0639**	-0.0208**
	(0.0194)	(0.0089)	(0.0186)	(0.0072)	(0.0484)	(0.0151)	(0.0119)	(0.0066)	(0.0402)	(0.0143)	(0.0048)	(0.0041)	(0.0155)	(0.0087)	(0.0287)	(0.0084)
Robust	-0.0403*	0.0079	-0.0442**	-0.0256***	-0.0687	0.0022	-0.0411***	-0.0182**	-0.1084**	-0.0105	-0.0085	-0.0107**	-0.0382**	-0.0076	-0.0639**	-0.0208**
	(0.0228)	(0.0108)	(0.0195)	(0.0084)	(0.0550)	(0.0181)	(0.0139)	(0.0076)	(0.0458)	(0.0161)	(0.0063)	(0.0048)	(0.0178)	(0.0100)	(0.0322)	(0.0094)
Mean DV	0.027	0.011	0.028	0.011	0.034	0.015	0.025	0.010	0.054	0.019	0.016	0.008	0.031	0.015	0.022	0.005
Standard Deviation	1.156	0.208	1.079	0.229	1.437	0.268	0.985	0.202	1.869	0.325	0.496	0.151	0.954	0.231	1.313	0.202
Observations within buffer	1877	3857	2720	2553	1273	2210	3128	3683	1605	3325	2945	3046	2331	3810	1828	2098
Observations untreated	16102	7492	15009	20661	10716	5092	22476	27036	13580	9346	8385	20683	17560	13881	19543	15822
Bandwidth untreated (km)	67.118	30.000	49.788	68.818	83.155	37.708	54.495	66.045	79.221	51.911	22.665	56.679	53.362	41.634	90.312	73.044
Bandwidth treated (km)	7.074	14.561	8.660	8.134	8.829	15.297	7.261	8.442	7.887	16.044	7.915	8.148	6.446	10.470	8.465	9.790

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 only. Cells that intersect with the limit of the 50km buffer of intervention are dropped from the sample. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Robustness of the RD results In the appendix, we present RD plots for a wider range of outcomes in figures A 3 and A 4.[18] Table A5 offers a detailed comparison of alternative coding approaches for the main violence outcomes. It includes results where the fatality numbers are subject to an inverse hyperbolic sine transformation. It also considers measures that include violence events that are coded less precisely (level 2 in ACLED). Including these less precisely coded violence events tends to makes the RD estimates less precise too. However, the broad patterns we found in our main results are generally robust to these alternative measurement approaches. Table A2 replicates the main RD results over the Western and Easter operations zones and finds no discontinuity in any outcome, which is consistent with the lower treatment intensity in these areas.

Table A3 further shows that the results are also robust to dropping cells crossed by the 50km buffer line, top-coding the 1% top fatalities and transforming the fatalities variable to a dummy indicating any fatality. Table reproduces the RDD results with manually set banwdith and a standard linear estimation of the G5 area effect while controlling for several transformation of the distance to the border and flexible controls for longitude and latitude, with stable coefficients across specifications.

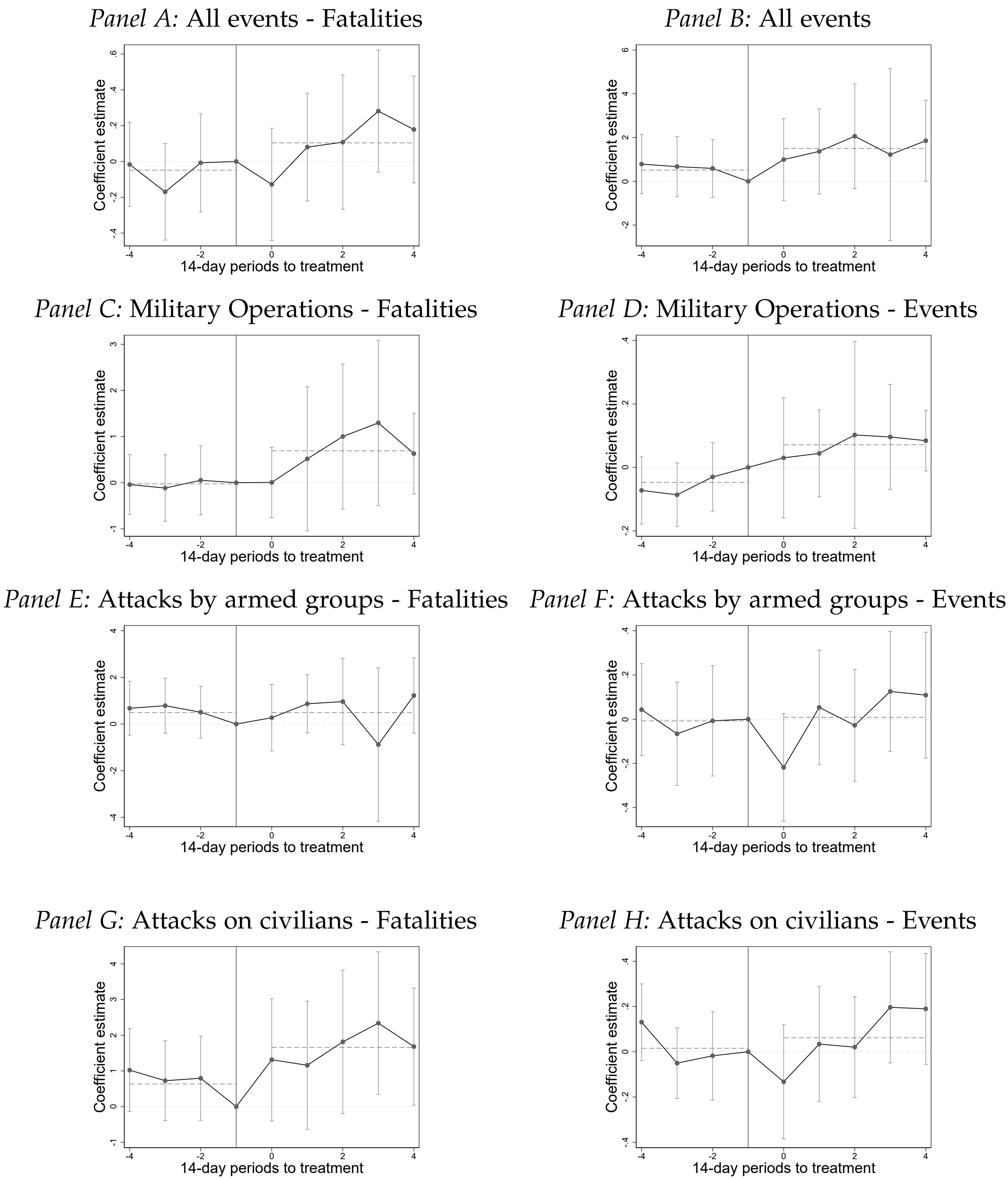
Strategic relocation of armed groups outside the G5 intervention zone. Despite the limitations of the RD setting preventing a direct assessment of gradual displacement beyond border areas, the observed patterns in the RDD plots do not support a symmetrical shift of conflict intensity just beyond the 50km operational zone of the G5 Joint Force. The concentration of violence in border regions, as evidenced by a significant linear increase in conflict intensity with proximity to borders, suggests that borders inherently escalate conflict intensity. Thus, relocating armed groups away from border areas should yield a positive impact on overall security.

Furthermore, the heterogeneity analysis reveals that the G5 intervention is most ef-

[18]One issue ”commongroup ==0” by the graphs in figure A 3 is the small number of security operations against Islamist groups in the optimal RD bandwidth. Figure A 4 shows this number increases when including less precisely coded events.Table A5 provides results for coarser geographical precision levels.

fective in regions with more porous borders. This indicates that the Joint Force has been successful in mitigating the border effect on conflict intensity. The second empirical exercise analysing response to French trigger operations in border areas highlights one mechanism through which this reduction operated.

Figure 4: Reaction to trigger events - border areas under G5



Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2020). Results are based on the estimating equation (2), in which we estimate time-to-treatment effects around the trigger operation for the interaction term $Border_i * G5_t$, and we include operation by time-to-treatment fixed effects. Standard errors are clustered at the grid-cell level, and grey bars represent 95% confidence intervals.

Table 4: Reaction to trigger events

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border x Post	-0.53	-0.15**	-0.51**	-0.10**	-0.02	-0.05	-0.68***	-0.07
	(0.40)	(0.07)	(0.21)	(0.04)	(0.36)	(0.06)	(0.21)	(0.04)
Border x Post x G5	0.99*	0.15*	0.99***	0.14***	-	0.02	1.03***	0.05
	(0.52)	(0.08)	(0.32)	(0.05)	(0.42)	(0.07)	(0.34)	(0.05)
Mean DV	0.693	0.178	0.313	0.054	0.380	0.125	0.360	0.099
Standard Deviation	4.676	0.576	3.373	0.372	2.635	0.387	2.432	0.332
Observations	3168	3168	3168	3168	3168	3168	3168	3168
Clusters	352	352	352	352	352	352	352	352

Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2020). Results are based on estimating equation (2). Standard errors are clustered at the grid-cell level and presented in parentheses; stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6.2 Response to trigger events

We now turn to our second empirical exercise, which compares responses of violence events to trigger operations. Figure 4 shows the differential violence in border areas when the G5 mission is active, split up in 2-week time periods around each trigger event. Panel A shows that relatively more violence events occur in border areas (so, within the G5-Sahel operation zone) when the G5-Sahel force is active. The split-up by type of violence in Panels B to D suggests that it is mostly the violence initiated by security forces that is driving this intensification. Focusing on panel B alone, this intensification is visible relatively soon after the trigger event. Table 4 confirms these patterns. Interestingly, it also shows that when the G5 mission is not active, there are less security operations following a trigger event in border areas.[19] This finding supports the hypothesis that security forces are hampered in their operations when the G5-Sahel mission is not active. However, this relative reduction in the intensity of operations is entirely off-set when the G5 mission is active. This result suggests that the G5-Sahel mission did achieve its goal of facilitating operations in border areas. This pattern is mirrored by violence against civilians. Additional analysis in table A7 suggests that this effect is mostly coming from violence by security forces against the civilian population.[20] In the main results, based on the RD approach, we found reductions in violence by militant groups. The response analysis does not show such reductions, but it is important to keep in mind that the response estimates are not set up to capture the longer-term dynamic impacts of security operations in border areas. For such global effects, we think our earlier RD results are

[19]Figure A 7 presents event studies around trigger events, comparing border and non-border areas, and showing these patterns separately for periods in which the G5-Sahel mission is active or not. Panel C shows the relative reduction in security operations in border areas after trigger events, and Panel D shows how this pattern reverses when the G5-Sahel mission is active.

[20]In this table, military operations against Islamist groups and communal militia respond in a similar way, which may appear at odds with the results of the RD analysis, where security force attacks against militia groups declined more than against Islamist groups. It should be kept in mind that the trigger events always involve Islamist groups, so that the nature of security force operations against communal militias may be different in trigger analysis. Indeed, it is impossible to disentangle the ethnic and religious dimensions of the conflict fully. In addition, it is also worth keeping in mind that there are limitations to the coding of actors (a substantial share of events involves unknown actors).

more insightful.

Robustness of the trigger event analysis In the appendix, we test the sensitivity of the results to an alternative criterion for the trigger event, where we focus on operations in which more than 10 (instead of at least 5) people died (Table A8). These results are noisier, as the more selective criterion reduces the sample, but the pattern on military operations is the same and significant (at 10%) for the event measures. In Table A9, we present results for an alternative coding of the G5 operation period, treating the entire period from September 2017 onwards as "active". In this coding, we ignore the incapacitation of the G5 mission after the 2018 suicide attack. Compared to table 4, we find similar patterns for military operations - they are lower in border areas before the G5 is active, and increase afterwards.[21] In table A10, we show a version of the response analysis where we include all grid cells in the sample, including those that were included in earlier event windows. The main results are robust to using this larger sample. Finally, table A11 shows the response analysis for binary violence outcomes. These findings confirm that the G5 mission changes response patterns on the extensive margin.

[21]In contrast to the main results, the coefficient on attacks against civilians is significantly positive before the G5 is active in the alternative coding. Low levels of violence before September 2017 hamper the comparability of violence patterns between the pre- and post-periods, so we do not want to emphasize this result. However, it is possible that the nature of violence against civilians changes over time. Before 2017, they could have suffered from poor security in border areas, whereas the intensification of the conflict might have them more vulnerable during military operations.

7 Conclusion

This paper examines how the establishment of an international armed force capable of crossing borders, known as the G5-Sahel, influenced the intensity and spatial distribution of conflict in the region's porous border areas. Our analysis indicates that the G5 mission reduced the intensity of conflict, at least locally, within its operation zone. By studying how the spatial distribution of violence responds to trigger events, we also find that the mission facilitated security operations in border areas. In this sense, our results offer a coherent narrative, whereby improved cooperation between national armies contributes to a reduction in equilibrium levels of violence. It should be kept in mind that the armies that are part of the G5-Sahel mission are regularly accused of human rights abuses, and we see that fatal violence against civilians mirrors their activities in our data. Hence, the welfare implications of our findings are far from clear in the context we study. In addition, the local effects that we estimate do not allow us to evaluate whether the G5 mission has helped to reduce levels of conflict at the aggregate level. The geographical spread of jihadist groups and the ongoing security challenges in the Sahel region put the local improvements we document in sharp perspective. In spite of these limitations, we think it is important to document that establishing zones in which national armies share security responsibilities can change conflict dynamics. These findings are particularly relevant for the many border regions in which armed groups exploit coordination frictions between national security forces.

References

Acemoglu, D., L. Fergusson, J. Robinson, D. Romero, and J. F. Vargas (2020). The Perils of High-Powered Incentives: Evidence from Colombia's False Positives. *American Economic Journal: Economic Policy 12*(3), 1–43.

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2015). Radio and the Rise of The Nazis in Prewar Germany *. *The Quarterly Journal of Economics 130*(4), 1885–1939.

Ager, P., L. Bursztyn, L. Leucht, and H.-J. Voth (2022). Killer Incentives: Rivalry, Performance and Risk-Taking among German Fighter Pilots, 1939–45. *The Review of Economic Studies 89*(5), 2257–2292.

Armand, A., P. Atwell, and J. F. Gomes (2020). The Reach of Radio: Ending Civil Conflict through Rebel Demobilization. *American Economic Review 110*(5), 1395–1429.

Berman, E., J. N. Shapiro, and J. H. Felter (2011). Can hearts and minds be bought? The economics of counterinsurgency in Iraq. *Journal of Political Economy 119*(4), 766 – 819.

Berman, N., M. Couttenier, D. Rohner, and M. Thoenig (2015). This Mine is Mine! How minerals fuel conflicts in Africa. *CESifo Working Paper Series* (July 2014).

Besley, T. and T. Persson (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton, NJ: Princeton University Press.

Blair, C. W. (2023). The Fortification Dilemma: Border Control and Rebel Violence *. *American Journal of Political Science*, forthcoming.

Boeke, S. and G. Chauzal (2017). BOEKE, Sergei; CHAUZAL,Grégory. *Annuaire français de relations internationales XVIII*.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica 82*(6), 2295–2326.

Calvo, T., E. Lavallée, M. Razafindrakoto, and F. Roubaud (2020). Fear Not For Man? Armed conflict and social capital in Mali. *Journal of Comparative Economics 48*(2), 251–276.

Cil, D., H. Fjelde, L. Hultman, and D. Nilsson (2020). Mapping blue helmets: Introducing the Geocoded Peacekeeping Operations (Geo-PKO) dataset. *Journal of Peace Research 57*(2).

Crost, B. and P. Johnston (2014). Aid under fire: Development projects and civil conflict. *American Economic Review 104*(6), 1833–1856.

Cunningham, D. E., K. S. Gleditsch, and I. Salehyan (2013). Non-state actors in civil wars: A new dataset. *Conflict Management and Peace Science 30*(5), 515–531.

de Chaisemartin, C. and X. D'Haultfoeuille (2022). Difference-in-Differences Estimators of Intertemporal Treatment Effects. *SSRN Electronic Journal*.

Dell, M. and P. Querubin (2018). Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies. *The Quarterly Journal of Economics 133*(2), 701–764.

Desmet, K., J. F. Gomes, and I. Ortuño-Ortín (2020). The geography of linguistic diversity and the provision of public goods. *Journal of Development Economics 143*.

Dincecco, M., J. Fenske, A. Menon, and S. Mukherjee (2022). Pre-Colonial Warfare and Long-Run Development in India. *The Economic Journal 132*(643), 981–1010.

Dube, O. and J. Vargas (2013). Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *The Review of Economic Studies*.

Durante, R. and E. Zhuravskaya (2018). Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict.

Eberle, U., D. Rohner, and M. Thoenig (2020). Heat and Hate : Climate Security and Farmer-Herder Conflicts in Africa. *ESOC Working Paper* (22), 79.

Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and Conflict: An Empirical Study. *American Economic Review 102*(4), 1310–1342.

Ferrara, E. L. and M. Harari (2018). Conflict, Climate, and Cells: A Disaggregated Analysis. *The Review of Economics and Statistics 100*(4), 594–608.

Fetzer, T. (2020). Can Workfare Programs Moderate Conflict? Evidence from India. *Journal of the European Economic Association*.

Fetzer, T. and S. Kyburz (2023). Does Democracy Moderate Resource-induced Conflicts ? *Review of Economics and Statistics*.

Fetzer, T., P. C. L. Souza, O. Vanden Eynde, and A. L. Wright (2021). Security transitions. *American Economic Review 111*(7).

Fetzer, T. R., O. Vanden Eynde, and A. L. Wright (2024). Team production on the battlefield : Evidence from NATO in Afghanistan. *PSE Working Paper*.

Fjelde, H., L. Hultman, and D. Nilsson (2019). Protection Through Presence: UN Peacekeeping and the Costs of Targeting Civilians. *International Organization 73*(1), 103–131.

Hultman, L., J. Kathman, and M. Shannon (2014). Beyond Keeping Peace: United Nations Effectiveness in the Midst of Fighting. *The American Political Science Review 108*(4), 737–753.

Martínez, L. R. (2017). Transnational insurgents: Evidence from Colombia's FARC at the border with Chávez's Venezuela. *Journal of Development Economics 126*, 138–153.

McGuirk, E. and N. Nunn (2022). Transhumant Pastoralism, Climate Change, and Conflict in Africa. *Working paper*.

Miguel, E., S. Satyanath, and E. Sergenti (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy 112*(4), 725–753.

Montalvo, J. and M. Reynal-Querol (2005). Ethnic polarization, potential conflict, and civil wars. *American Economic Review 95*(3), 796–816.

Mueller, H., D. Rohner, and D. Schönholzer (2022). Ethnic Violence Across Space. *The Economic Journal 132*(642), 709–740.

Premand, P. and D. Rohner (2023). Cash and Conflict Large-Scale Experimental Evidence from Niger. *Working paper*.

Richard, M. (2022). An Indirect Cost of Conflict : Insecurity as a Barrier to Seasonal Migration in Mali. *Working paper*.

Rohner, D. (2024). Mediation, Military, and Money: The Promises and Pitfalls of Outside Interventions to End Armed Conflicts. *Journal of Economic Literature 62*(1), 155–195.

Sanchez de la Sierra, R. (2020). On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo. *Journal of Political Economy*.

Shapiro, J. N. and O. Vanden Eynde (2023). Fiscal Incentives for Conflict: Evidence from India's Red Corridor. *The Review of Economics and Statistics 105*(1), 217–225.

Tai, X. H., S. Mehra, and J. E. Blumenstock (2022). Mobile phone data reveal the effects of violence on internal displacement in Afghanistan. *Nature Human Behaviour 6*(5).

Tapsoba, A. (2023). The Cost of Fear: Impact of Violence Risk on Child Health During Conflict.

Touchard, L. (2018). Quels défis pour la Force. *Areion Group 134*(134), 44–49.

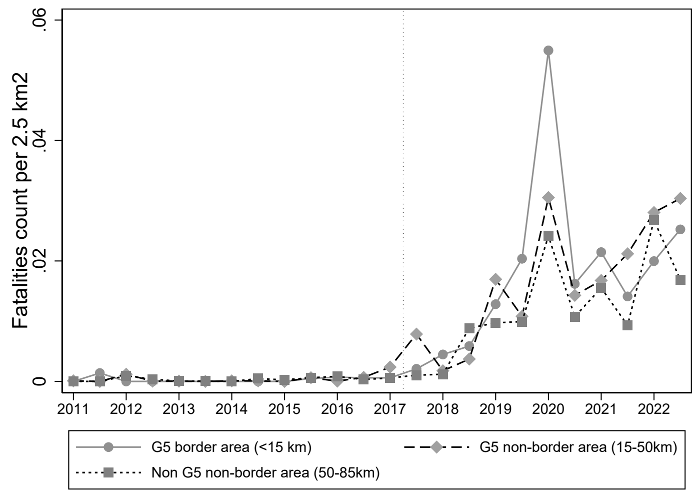
Vanden Eynde, O. (2016). Targets of Violence: Evidence from India's Naxalite Conflict. *The Economic Journal* (November), 1–50.

Yanagizawa-Drott, D. (2014). Propaganda and Conflict: Evidence from the Rwandan Genocide *. *The Quarterly Journal of Economics 129*(4), 1947–1994.

Appendix to **Cooperation between National Armies**

For Online Publication

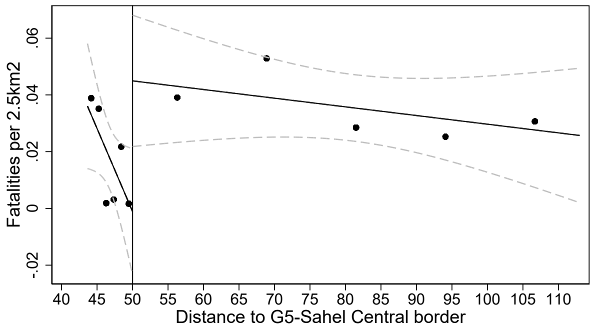
Figure A 1: Time trends in conflict intensity in G5 Sahel central operation zone



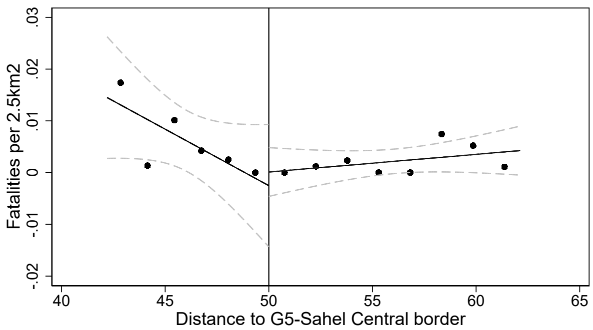
Notes: Observations at the grid-cell level, binned in six months periods (2011-2022). The sample comprises the central Zone (Mali-Niger-Burkina Faso three borders regions). Included events coded with geo-precision level 1 only.

Figure A 2: Regression Discontinuity for G5 operation zone

Panel A: All events (after Sept 2017)

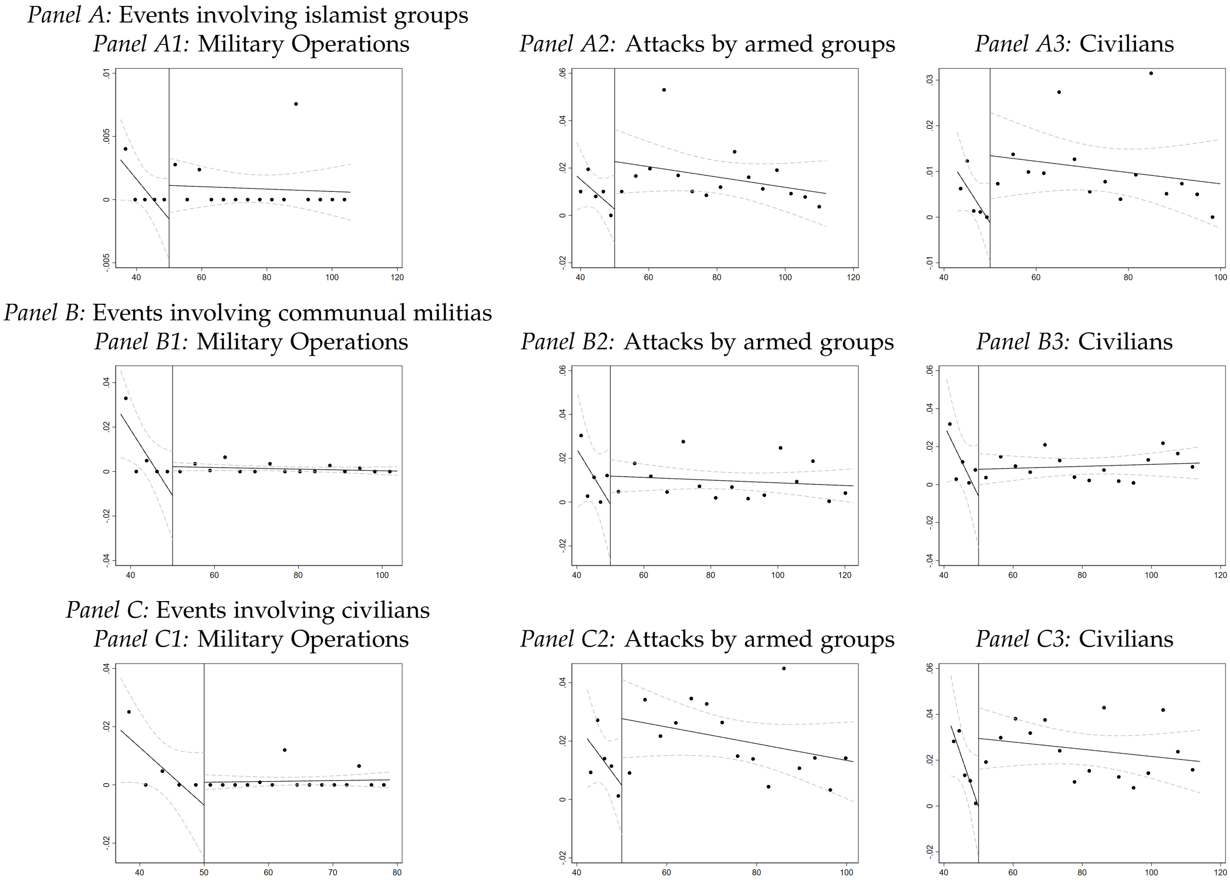


Panel B: All events (before Sept 2017)



Notes: Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 200km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. The sample comprises the central Zone (Mali-Niger-Burkina Faso three borders regions). Included events are coded with geo-precision level 1.

Figure A 3: RDD plots with optimal buffer for sub-groups, geo-precision level 1



Notes: Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. All conflict events occurring for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 only.

Table A1: Discontinuity in conflict intensity for central G5-Sahel operation zones 2017-2020, by country/border.

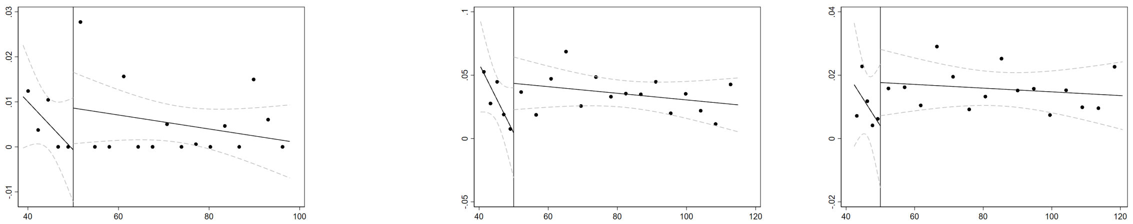
	Mali		Burkina Faso		Niger		MLI-BFA		MLI-NER		NER-BFA	
	Fatalities	Events	Fatalities	Events	Fatalities	Events	Fatalities	Events	Fatalities	Events	Fatalities	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Conventional	-0.0399*	-0.0088	-0.0842**	-0.0246**	-0.0042***	-0.0052***	-0.0813*	-0.0124	-0.0198***	-0.0056	-0.0073	-0.0064
	(0.0213)	(0.0109)	(0.0369)	(0.0114)	(0.0016)	(0.0017)	(0.0460)	(0.0119)	(0.0066)	(0.0053)	(0.0104)	(0.0077)
Bias-corrected	-0.0499**	-0.0081	-0.0802**	-0.0271**	-0.0042***	-0.0054***	-0.1031**	-0.0169	-0.0150**	-0.0063	-0.0029	-0.0066
	(0.0213)	(0.0109)	(0.0369)	(0.0114)	(0.0016)	(0.0017)	(0.0460)	(0.0119)	(0.0066)	(0.0053)	(0.0104)	(0.0077)
Robust	-0.0499**	-0.0081	-0.0802**	-0.0271*	-0.0042*	-0.0054**	-0.1031*	-0.0169	-0.0150**	-0.0063	-0.0029	-0.0066
	(0.0246)	(0.0125)	(0.0382)	(0.0139)	(0.0021)	(0.0022)	(0.0613)	(0.0141)	(0.0074)	(0.0065)	(0.0109)	(0.0083)
Mean DV	0.026	0.011	0.045	0.020	0.013	0.004	0.033	0.014	0.021	0.008	0.028	0.012
Standard Deviation	0.999	0.243	1.213	0.270	1.197	0.085	1.081	0.253	1.077	0.199	1.205	0.198
Observations within buffer	1592	4393	1736	1489	950	1475	2852	3114	7180	3559	3871	3364
Observations untreated	16791	14490	8179	7730	11640	9115	35848	27757	30602	38130	6194	14467
Bandwidth untreated (km)	72.991	63.214	52.278	49.213	73.416	57.273	130.390	102.342	126.416	154.848	49.630	115.350
Bandwidth treated (km)	6.800	18.233	9.986	8.625	5.728	8.952	10.506	11.491	31.853	15.773	30.229	26.307

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 only. Cells that intersect with the limit of the 50km buffer of intervention are dropped from the sample. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** p<0.01, ** p<0.05, * p<0.1.

Figure A 4: RDD plots with optimal buffer for sub-groups, geo-precision level 1 and 2

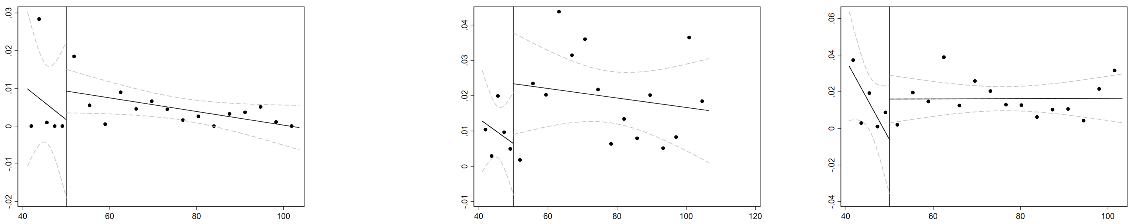
Panel A: Events involving islamist groups

Panel A1: Military Operations *Panel A2:* Attacks by armed groups *Panel A3:* Civilians



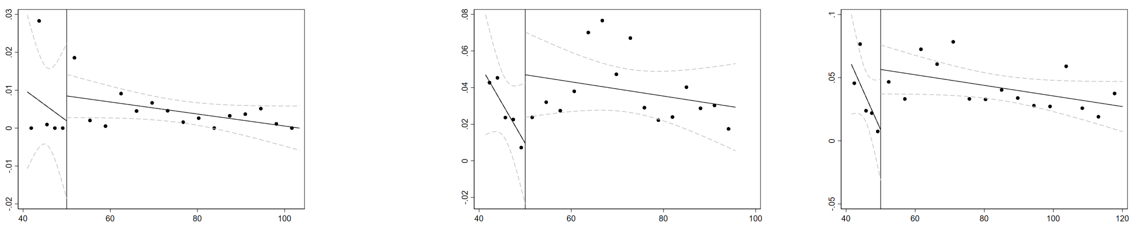
Panel B: Events involving communual militias

Panel B1: Military Operations *Panel B2:* Attacks by armed groups *Panel B3:* Civilians



Panel C: Events involving civilians

Panel C1: Military Operations *Panel C2:* Attacks by armed groups *Panel C3:* Civilians



Notes: Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. All conflict events occurring for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 and 2.

Table A2: Discontinuity in conflict intensity for all G5-Sahel operation zones 2017-2020, all regions.

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Central Zone (Mali-Niger-Burkina Faso three borders regions)								
Robust	-0.0452***	-0.0092	-0.0045*	-0.0022***	-0.0356**	-0.0056	-0.0321***	-0.0024
	(0.0161)	(0.0067)	(0.0024)	(0.0007)	(0.0143)	(0.0063)	(0.0114)	(0.0047)
Mean DV	0.026	0.010	0.003	0.001	0.022	0.009	0.015	0.007
Standard Deviation	1.071	0.200	0.211	0.032	1.030	0.178	0.640	0.119
Observations within buffer	3641	5821	7214	4863	3570	6358	3201	7491
Observations untreated	34378	21575	29292	44030	29242	19120	29838	18975
Bandwidth untreated (km)	63.484	39.229	53.794	81.883	53.718	34.596	54.787	34.321
Bandwidth treated (km)	6.347	10.099	12.443	8.409	6.189	10.973	5.569	12.931
Panel B: Eastern Zone (Niger-Chad border)								
Robust	-0.0138	-0.0003	-0.0018	-0.0003	-0.0108	-0.0005	-0.0035	0.0011
	(0.0148)	(0.0045)	(0.0016)	(0.0002)	(0.0120)	(0.0043)	(0.0044)	(0.0045)
Mean DV	0.004	0.002	0.001	0.000	0.003	0.002	0.002	0.002
Standard Deviation	0.332	0.098	0.142	0.015	0.256	0.085	0.137	0.080
Observations within buffer	3904	8375	3587	3587	3935	8259	3927	8598
Observations untreated	13977	24559	9555	8462	18746	22374	9858	17576
Bandwidth untreated (km)	45.304	78.924	30.996	27.498	60.539	72.039	32.022	56.787
Bandwidth treated (km)	12.819	27.554	11.753	11.753	12.904	27.168	12.884	28.231
Panel C: Western Zone (Mali-Mauritania border)								
Robust		-0.0002			-0.0003		-0.0002	
		(0.0003)			(0.0003)		(0.0002)	
Mean DV		0.002			0.001		0.001	
Standard Deviation		0.058			0.055		0.038	
Observations within buffer		2750			2750		2589	
Observations untreated		11367			13264		14096	
Bandwidth untreated (km)		20.655			24.049		25.824	
Bandwidth treated (km)		4.771			4.771		4.521	

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 only. Empty cells means the discontinuity could not be estimated due to lack of variability in the dependent variable. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Discontinuity in conflict intensity for central G5-Sahel operation zones 2017-2020, Robustness checks.

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Splitted cells dropped								
Robust	-0.0525**	-0.0284***	-0.0023	-0.0031***	-0.0449**	-0.0232***	-0.0231	-0.0144***
	(0.0237)	(0.0090)	(0.0048)	(0.0009)	(0.0224)	(0.0082)	(0.0148)	(0.0047)
Mean DV	0.028	0.011	0.003	0.001	0.024	0.010	0.016	0.007
Standard Deviation	1.122	0.220	0.213	0.033	1.080	0.198	0.648	0.123
Observations within buffer	1930	3069	2646	4005	1833	3280	1463	4178
Observations untreated	42957	38014	41047	49234	37312	32178	34938	33698
Bandwidth untreated (km)	81.727	72.396	78.116	93.936	70.949	61.274	66.506	64.139
Bandwidth treated (km)	4.944	6.966	6.113	8.479	4.765	7.291	4.218	8.756
Panel B: Top-coded fatalities (0.01)								
Robust	-0.0394***	-0.0110	-0.0054**	-0.0022***	-0.0315**	-0.0074	-0.0320***	-0.0034
	(0.0144)	(0.0073)	(0.0025)	(0.0007)	(0.0130)	(0.0069)	(0.0117)	(0.0051)
Mean DV	0.024	0.011	0.003	0.001	0.020	0.010	0.014	0.007
Standard Deviation	0.748	0.220	0.203	0.033	0.687	0.198	0.474	0.123
Observations within buffer	5131	5878	7379	4865	5089	6358	4621	7606
Observations untreated	35693	25267	27394	43013	29074	22521	33368	21791
Bandwidth untreated (km)	65.999	46.142	50.227	80.011	53.424	40.894	61.575	39.637
Bandwidth treated (km)	8.874	10.204	12.752	8.414	8.789	10.976	7.997	13.119
Panel C: Any fatality								
Robust	-0.0049**	-0.0110	-0.0010***	-0.0022***	-0.0042*	-0.0074	-0.0042**	-0.0034
	(0.0023)	(0.0073)	(0.0004)	(0.0007)	(0.0023)	(0.0069)	(0.0020)	(0.0051)
Mean DV	0.004	0.011	0.000	0.001	0.003	0.010	0.003	0.007
Standard Deviation	0.060	0.220	0.020	0.033	0.058	0.198	0.051	0.123
Observations within buffer	5586	5878	6355	4865	5613	6358	5166	7606
Observations untreated	34557	25267	41484	43013	31794	22521	32913	21791
Bandwidth untreated (km)	63.862	46.142	77.063	80.011	58.605	40.894	60.730	39.637
Bandwidth treated (km)	9.714	10.204	10.965	8.414	9.737	10.976	8.955	13.119

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 only. Cells that intersect with the limit of the 50km buffer of intervention are dropped from the sample. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Discontinuity in conflict intensity for central G5-Sahel operation zones 2017-2020, OLS estimations with flexible controls for coordinates.

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 15-100 km buffer								
Linear distance ctrl	-0.0297**	-0.0053	0.0023	-0.0005	-0.0297**	-0.0049	-0.0162*	-0.0010
	(0.0137)	(0.0043)	(0.0034)	(0.0006)	(0.0128)	(0.0039)	(0.0090)	(0.0023)
Linear lat. lon. ctrl	-0.1247*	-0.0105	-0.0170	-0.0055*	-0.1078	-0.0035	-0.0616	-0.0029
	(0.0740)	(0.0233)	(0.0182)	(0.0030)	(0.0691)	(0.0211)	(0.0486)	(0.0127)
Linear dist. lat. lon. ctrl	-0.0301**	-0.0055	0.0022	-0.0005	-0.0301**	-0.0051	-0.0165*	-0.0011
	(0.0137)	(0.0043)	(0.0034)	(0.0006)	(0.0128)	(0.0039)	(0.0090)	(0.0023)
Linear dist. sq lat. lon. ctrl	-0.0295**	-0.0053	0.0023	-0.0005	-0.0296**	-0.0049	-0.0161*	-0.0010
	(0.0137)	(0.0043)	(0.0034)	(0.0006)	(0.0128)	(0.0039)	(0.0090)	(0.0023)
Linear dist. lat. lon. & flex. lat. lon. ctrl	-0.0296**	-0.0053	0.0023	-0.0005	-0.0297**	-0.0049	-0.0162*	-0.0010
	(0.0137)	(0.0043)	(0.0034)	(0.0006)	(0.0128)	(0.0039)	(0.0090)	(0.0023)
Linear, sq and cubic distance ctrl & flex. lat. lon. ctrl	-0.0278	-0.0081	-0.0077	-0.0018	-0.0134	-0.0052	-0.0176	0.0000
	(0.0274)	(0.0086)	(0.0068)	(0.0011)	(0.0256)	(0.0078)	(0.0180)	(0.0047)
Mean DV								
Standard Deviation								
Observations within buffer	74332	74332	74332	74332	74332	74332	74332	74332
Panel B: 50 km buffer								
Linear distance ctrl	-0.0266**	-0.0034	-0.0002	-0.0006	-0.0241**	-0.0027	-0.0227***	-0.0006
	(0.0129)	(0.0037)	(0.0035)	(0.0005)	(0.0119)	(0.0034)	(0.0082)	(0.0021)
Linear lat. lon. ctrl	-0.1259*	-0.0049	-0.0258	-0.0062**	-0.1010	0.0019	-0.0738*	-0.0011
	(0.0683)	(0.0198)	(0.0186)	(0.0029)	(0.0631)	(0.0178)	(0.0436)	(0.0110)
Linear dist. lat. lon. ctrl	-0.0274**	-0.0036	-0.0003	-0.0006	-0.0247**	-0.0029	-0.0232***	-0.0007
	(0.0129)	(0.0037)	(0.0035)	(0.0005)	(0.0119)	(0.0034)	(0.0082)	(0.0021)
Linear dist. sq lat. lon. ctrl	-0.0265**	-0.0033	-0.0002	-0.0005	-0.0240**	-0.0026	-0.0226***	-0.0005
	(0.0129)	(0.0037)	(0.0035)	(0.0005)	(0.0119)	(0.0034)	(0.0082)	(0.0021)
Linear dist. lat. lon. & flex. lat. lon. ctrl	-0.0268**	-0.0034	-0.0002	-0.0006	-0.0242**	-0.0027	-0.0229***	-0.0006
	(0.0129)	(0.0037)	(0.0035)	(0.0005)	(0.0119)	(0.0034)	(0.0082)	(0.0021)
Linear, sq and cubic distance ctrl & flex. lat. lon. ctrl	-0.0240	-0.0061	-0.0047	-0.0016	-0.0139	-0.0038	-0.0120	-0.0001
	(0.0258)	(0.0075)	(0.0070)	(0.0011)	(0.0239)	(0.0067)	(0.0165)	(0.0041)
Mean DV								
Standard Deviation								
Observations within buffer	84395	84395	84395	84395	84395	84395	84395	84395

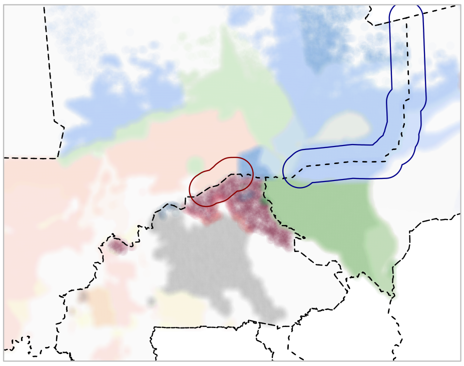
Notes : Manual choice of bandwidths. Additional country controls, no border segment controls. All conflict events occurring within 250km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. Include events coded with geo-precision level 1 only. Only estimate of coefficient of interest "Cell within G5 operation zone" reported. Control variables include distance to the G5 central border, latitude, longitude, and squared and cubic transformations of latitude and longitude. Distance to the border and its transformation and are allowed a different slope within and outside of the operation zone while the effect of latitude and longitude is assumed to be the same within and outside of the G5 area to avoid overfitting. Robust standard errors in parentheses- *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Discontinuity in conflict intensity for G5 Sahel Central operation zone 2017-2020 , robustness checks.

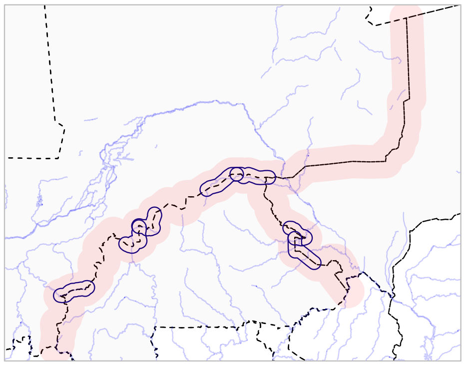
	Military operations						Attacks by armed groups						Attacks on Civilians					
	Geo1			Geo2			Geo1			Geo2			Geo1			Geo2		
	Events	Fatalities(#)	Fatalities(IHS)	Events	Fatalities(#)	Fatalities(IHS)	Events	Fatalities(#)	Fatalities(IHS)	Events	Fatalities(#)	Fatalities(IHS)	Events	Fatalities(#)	Fatalities(IHS)	Events	Fatalities(#)	Fatalities(IHS)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Panel 0: All Events																		
Conventional	-0.0020***	-0.0046**	-0.0018***	-0.0032**	-0.0280**	-0.0056***	-0.0054	-0.0349***	-0.0085**	-0.0013	-0.0194	-0.0025	-0.0019	-0.0299***	-0.0087***	0.0005	-0.0499***	-0.0081*
	(0.0006)	(0.0021)	(0.0007)	(0.0015)	(0.0119)	(0.0014)	(0.0055)	(0.0135)	(0.0034)	(0.0090)	(0.0174)	(0.0050)	(0.0041)	(0.0099)	(0.0030)	(0.0066)	(0.0180)	(0.0043)
Bias-corrected	-0.0022***	-0.0045**	-0.0021***	-0.0036**	-0.0355***	-0.0064***	-0.0056	-0.0356***	-0.0092***	-0.0015	-0.0187	-0.0023	-0.0024	-0.0321***	-0.0095***	-0.0002	-0.0530***	-0.0085**
	(0.0006)	(0.0021)	(0.0007)	(0.0015)	(0.0119)	(0.0014)	(0.0055)	(0.0135)	(0.0034)	(0.0090)	(0.0174)	(0.0050)	(0.0041)	(0.0099)	(0.0030)	(0.0066)	(0.0180)	(0.0043)
Robust	-0.0022***	-0.0045*	-0.0021**	-0.0036**	-0.0355**	-0.0064***	-0.0056	-0.0356**	-0.0092**	-0.0015	-0.0187	-0.0023	-0.0024	-0.0321***	-0.0095***	-0.0002	-0.0530***	-0.0085*
	(0.0007)	(0.0024)	(0.0008)	(0.0018)	(0.0151)	(0.0018)	(0.0063)	(0.0143)	(0.0040)	(0.0104)	(0.0192)	(0.0058)	(0.0047)	(0.0114)	(0.0036)	(0.0075)	(0.0206)	(0.0051)
Observations within buffer	4863	7214	5751	5789	5058	4798	6358	3570	5053	6298	4142	6011	7491	3201	4391	7558	3829	5759
Observations untreated	44030	29292	39076	33127	27298	42675	19120	29242	24499	14974	19095	17148	18975	29838	27214	15163	32646	24930
Bandwidth untreated (km)	81.883	53.794	72.495	61.152	50.043	79.308	34.596	53.718	44.758	26.884	34.538	30.886	34.321	54.787	49.873	27.252	60.108	45.535
Bandwidth treated (km)	8.409	12.443	9.997	10.051	8.738	8.282	10.973	6.189	8.726	10.868	7.238	10.403	12.931	5.569	7.665	13.044	6.734	10.013
Panel A: Events involving Islamist groups																		
Conventional	-0.0006	-0.0012	-0.0006	-0.0013	-0.0129	-0.0023**	-0.0083***	-0.0238**	-0.0077***	-0.0048	-0.0222*	-0.0040	-0.0035*	-0.0120***	-0.0058***	-0.0022	-0.0153*	-0.0035
	(0.0004)	(0.0012)	(0.0005)	(0.0010)	(0.0084)	(0.0010)	(0.0029)	(0.0104)	(0.0016)	(0.0063)	(0.0126)	(0.0035)	(0.0019)	(0.0040)	(0.0013)	(0.0035)	(0.0088)	(0.0025)
Bias-corrected	-0.0008*	-0.0014	-0.0007	-0.0015	-0.0162*	-0.0026***	-0.0086***	-0.0233**	-0.0082***	-0.0047	-0.0199	-0.0041	-0.0039**	-0.0127***	-0.0062***	-0.0024	-0.0165*	-0.0037
	(0.0004)	(0.0012)	(0.0005)	(0.0010)	(0.0084)	(0.0010)	(0.0029)	(0.0104)	(0.0016)	(0.0063)	(0.0126)	(0.0035)	(0.0019)	(0.0040)	(0.0013)	(0.0035)	(0.0088)	(0.0025)
Robust	-0.0008	-0.0014	-0.0007	-0.0015	-0.0162	-0.0026**	-0.0086**	-0.0233**	-0.0082***	-0.0047	-0.0199	-0.0041	-0.0039*	-0.0127***	-0.0062***	-0.0024	-0.0165*	-0.0037
	(0.0005)	(0.0016)	(0.0006)	(0.0013)	(0.0107)	(0.0013)	(0.0034)	(0.0110)	(0.0019)	(0.0074)	(0.0139)	(0.0041)	(0.0022)	(0.0045)	(0.0015)	(0.0041)	(0.0100)	(0.0030)
Observations within buffer	5862	8375	7966	8043	6400	5318	5216	6358	4362	6216	5476	6227	8805	4082	4305	10042	4433	7214
Observations untreated	29991	31396	33112	39346	26259	36853	31051	33063	31190	18758	24261	23732	31039	22979	34691	22650	33467	35520
Bandwidth untreated (km)	55.143	57.782	61.133	72.892	48.054	68.135	57.037	61.027	57.349	33.943	44.218	43.266	57.018	41.878	64.116	41.144	61.729	65.589
Bandwidth treated (km)	10.167	14.412	13.678	13.835	11.047	9.272	9.051	10.971	7.606	10.709	9.587	10.729	15.159	7.100	7.507	17.238	7.726	12.443
Panel B: Events involving Communal militia																		
Conventional	-0.0015***	-0.0024*	-0.0013**	-0.0022***	-0.0145*	-0.0029***	0.0018	-0.0071	-0.0007	0.0031	-0.0126	0.0012	0.0028	-0.0073	-0.0007	0.0042	-0.0096	0.0008
	(0.0005)	(0.0014)	(0.0005)	(0.0007)	(0.0080)	(0.0009)	(0.0028)	(0.0078)	(0.0022)	(0.0033)	(0.0091)	(0.0027)	(0.0024)	(0.0068)	(0.0021)	(0.0028)	(0.0079)	(0.0023)
Bias-corrected	-0.0017***	-0.0023*	-0.0016***	-0.0023***	-0.0191**	-0.0034***	0.0025	-0.0087	-0.0005	0.0036	-0.0134	0.0012	0.0036	-0.0086	-0.0010	0.0047	-0.0105	0.0007
	(0.0005)	(0.0014)	(0.0005)	(0.0007)	(0.0080)	(0.0009)	(0.0028)	(0.0078)	(0.0022)	(0.0033)	(0.0091)	(0.0027)	(0.0024)	(0.0068)	(0.0021)	(0.0028)	(0.0079)	(0.0023)
Robust	-0.0017***	-0.0023	-0.0016**	-0.0023***	-0.0191*	-0.0034***	0.0025	-0.0087	-0.0005	0.0036	-0.0134	0.0012	0.0036	-0.0086	-0.0010	0.0047	-0.0105	0.0007
	(0.0006)	(0.0018)	(0.0006)	(0.0008)	(0.0109)	(0.0012)	(0.0034)	(0.0088)	(0.0026)	(0.0039)	(0.0100)	(0.0031)	(0.0029)	(0.0078)	(0.0024)	(0.0034)	(0.0088)	(0.0026)
Observations within buffer	4755	6890	5476	5091	5229	5338	10870	3309	9967	10823	3303	8302	11883	3130	6965	10822	3246	6980
Observations untreated	34435	30027	38051	28846	28834	46875	20431	34722	27457	14629	25405	14983	19491	32549	23613	13150	24587	15319
Bandwidth untreated (km)	63.599	55.228	70.365	53.008	52.982	87.325	37.051	64.177	50.360	26.269	46.299	26.905	35.277	59.875	43.029	23.576	44.910	27.519
Bandwidth treated (km)	8.221	11.943	9.588	8.793	9.080	9.306	18.590	5.764	17.098	18.525	5.735	14.268	20.289	5.447	12.098	18.521	5.619	12.137
Panel C: Events involving civilians																		
Conventional	-0.0010***	-0.0023	-0.0007**	-0.0017***	-0.0133*	-0.0026***	0.0005	-0.0215**	-0.0065**	0.0032	-0.0226	-0.0035	-0.0019	-0.0299***	-0.0087***	0.0005	-0.0499***	-0.0081*
	(0.0003)	(0.0022)	(0.0004)	(0.0005)	(0.0080)	(0.0008)	(0.0037)	(0.0092)	(0.0029)	(0.0059)	(0.0151)	(0.0041)	(0.0041)	(0.0099)	(0.0030)	(0.0066)	(0.0180)	(0.0043)
Bias-corrected	-0.0011***	-0.0023	-0.0010***	-0.0018***	-0.0179**	-0.0031***	0.0005	-0.0232**	-0.0072**	0.0033	-0.0243	-0.0037	-0.0024	-0.0321***	-0.0095***	-0.0002	-0.0530***	-0.0085**
	(0.0003)	(0.0022)	(0.0004)	(0.0005)	(0.0080)	(0.0008)	(0.0037)	(0.0092)	(0.0029)	(0.0059)	(0.0151)	(0.0041)	(0.0041)	(0.0099)	(0.0030)	(0.0066)	(0.0180)	(0.0043)
Robust	-0.0011***	-0.0023	-0.0010**	-0.0018***	-0.0179	-0.0031***	0.0005	-0.0232**	-0.0072**	0.0033	-0.0243	-0.0037	-0.0024	-0.0321***	-0.0095***	-0.0002	-0.0530***	-0.0085*
	(0.0004)	(0.0015)	(0.0004)	(0.0007)	(0.0109)	(0.0011)	(0.0043)	(0.0102)	(0.0033)	(0.0068)	(0.0165)	(0.0047)	(0.0047)	(0.0114)	(0.0036)	(0.0075)	(0.0206)	(0.0051)
Observations within buffer	5133	7755	5965	5004	5258	5342	9893	3114	4497	9816	3758	6347	7491	3201	4391	7558	3829	5759
Observations untreated	35314	15956	25090	30217	28488	48181	17759	24679	23498	14268	20983	19806	18975	29838	27214	15163	32646	24930
Bandwidth untreated (km)	65.140	28.706	45.821	55.599	52.221	89.869	32.063	45.074	42.811	25.545	38.073	35.807	34.321	54.787	49.873	27.252	60.108	45.535
Bandwidth treated (km)	8.884	13.360	10.330	8.659	9.149	9.323	16.967	5.422	7.829	16.837	6.566	10.943	12.931	5.569	7.665	13.044	6.734	10.013

Notes : Discontinuity estimated at 50 km. Data-driven choice of two bandwidths for MSE-optimal point estimation. Local polynomial of order 1. Additional country and border controls. All conflict events occurring within 250km of each considered G5-Sahel borders are included for the period between September 2017 to January 2020. Military operations involving islamist groups record fatalities from islamist groups only, whereas attacks by armed group and attacks on civilians involving islamist goups record fatalities caused by islamist groups. Idem for communal militia and unidentified armed groups. Empty cells means the discontinuity could not be estimated due to lack of variability in the dependent variable. Robust Calonico-Cattaneo-Titiunik standard errors in parentheses- *** p<0.01, ** p<0.05, * p<0.1.

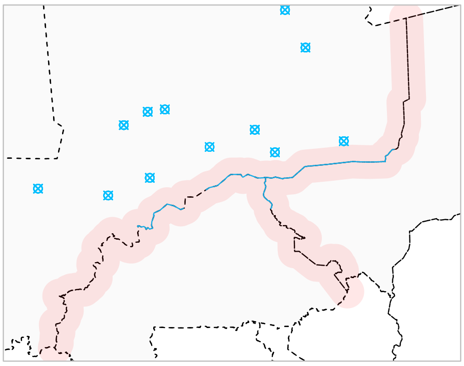
Figure A 5: Border segments characteristics



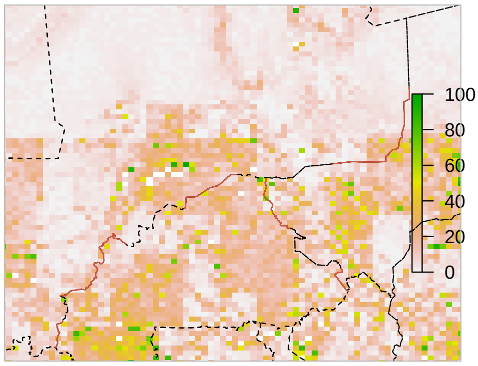
(a) Segments with transborder ethnic groups



(b) Segments following rivers



(c) Segments close to MINSUMA bases



(d) Segments in high rugged areas

Notes: Segments with transborder ethnic groups are those segments where the same ethnic group accounts for at least 10% of the population on both sides of the border. Such segments are represented in red for the Fulani group and in blue for the Tuareg. The other groups (Songhay in green, Mossi in gray, Bambara in yellow) never represents more than 10% of the population on both sides of the border. Segments aligned with rivers are defined as those segments within five kilometers of a river over 40% of their total length. Segments ”close” to MINSUMA are segments for which the minimum distance between a MINUSMA base and the segment is lower than the median distance. Segments in ”high rugged” areas are segments for which the average ruggedness of cells crossed by the segments is above the median ruggedness of cells crossed by the other segments.

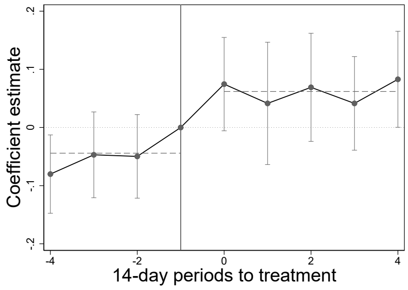
Table A6: Discontinuity in conflict intensity for central G5-Sahel operation zones 2017-2020, by country/border.

	Mali	Burkina Faso	Niger	MLI-BFA	MLI-NER	NER-BFA
	(1)	(2)	(3)	(4)	(5)	(6)
Share fulani	0.117	0.151	0.042	0.152	0.043	0.217
	(0.165)	(0.160)	(0.088)	(0.168)	(0.095)	(0.178)
Border segments with transborder fulani group	0.101	0.030	0.000	0.092	0.000	0.000
	(0.302)	(0.170)	(0.000)	(0.289)	(0.000)	(0.000)
Share tuareg	0.140	0.015	0.267	0.035	0.314	0.032
	(0.160)	(0.026)	(0.147)	(0.048)	(0.086)	(0.058)
Border segments with transborder tuareg group	0.322	0.000	0.763	0.000	0.835	0.000
	(0.467)	(0.000)	(0.425)	(0.000)	(0.371)	(0.000)
Border segments with any transborder ethnic group	0.424	0.030	0.763	0.092	0.835	0.000
	(0.494)	(0.170)	(0.425)	(0.289)	(0.371)	(0.000)
Distance to MINUSMA base	170.566	203.738	203.274	166.226	153.110	237.136
	(93.122)	(88.597)	(94.203)	(87.805)	(87.898)	(111.973)
Average rugdness around border segment	11.753	9.663	12.557	11.003	10.898	11.469
	(7.601)	(6.638)	(8.329)	(6.829)	(8.729)	(6.703)
Border segment aligns with river	0.270	0.439	0.057	0.384	0.039	0.328
	(0.444)	(0.496)	(0.231)	(0.486)	(0.192)	(0.470)
Observations	142898	71944	81068	29904	24034	15846
G5 interventions in acled	23	8	14	24	13	22

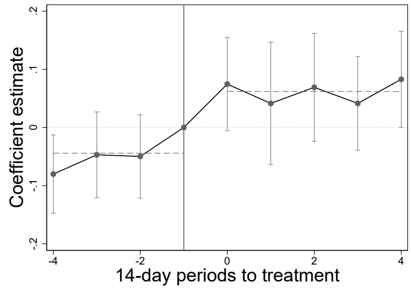
Notes: Observations at the grid-cell level in the first part of the table, and total count of interventions in the second part. Each grid-cell is defined by the closest border segment. Segments with transborder ethnic groups are those segments where the same ethnic group accounts for at least 10% of the population on both sides of the border. Segments aligned with rivers are defined as those segments within five km of a river over 40% of their total length. Segments in ”high rugged” areas are segments for which the average ruggedness of cells crossed by the segments is above the median ruggedness of cells crossed by the other segments.

Figure A 6: Reaction to trigger events - time patterns (no differences)

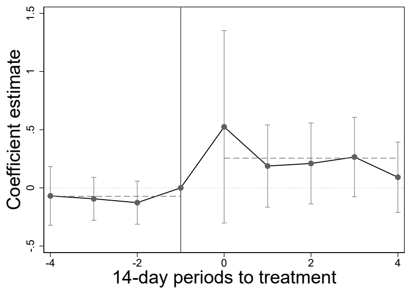
Panel A: All events - Fatalities



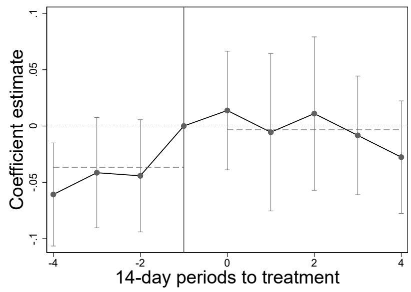
Panel B: All events



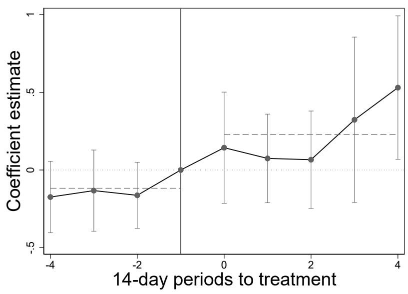
Panel C: Military Operations - Fatalities



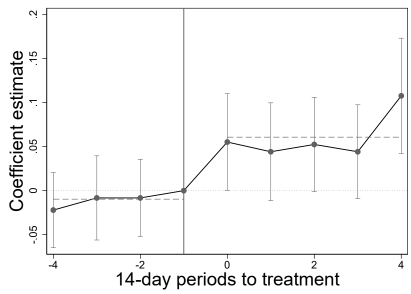
Panel D: Military Operations - Events



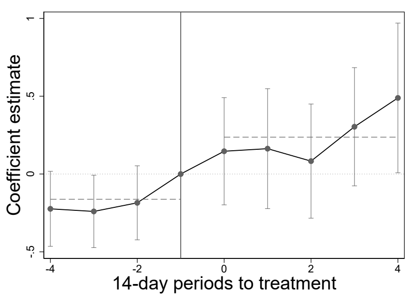
Panel E: Attacks by armed groups - Fatalities



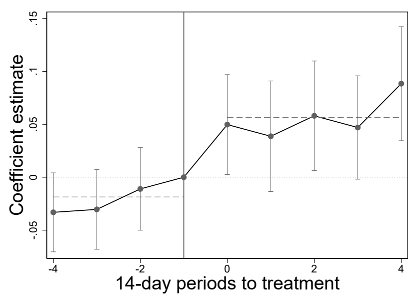
Panel F: Attacks by armed groups - Events



Panel G: Attacks by armed groups - Fatalities



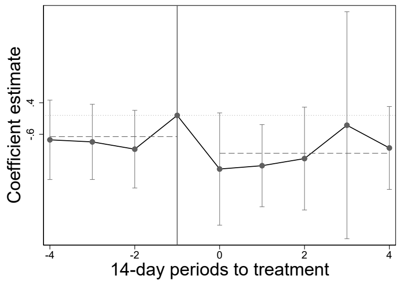
Panel H: Attacks by armed groups - Events



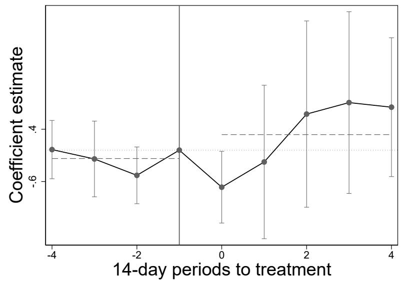
Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2021). The graph show coefficients on time-to-treatment dummies around the trigger operation. The model includes operation fixed effects. Standard errors are clustered at the grid-cell level, and grey bars represent 95% confidence intervals.

Figure A 7: Reaction to trigger events - comparison of border areas

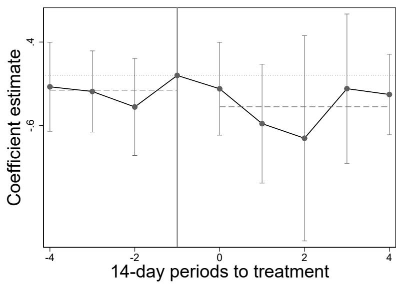
Panel A: All events -fatalities (no G5)



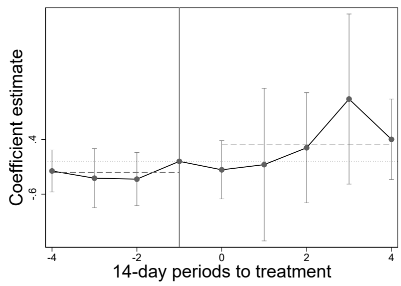
Panel B: All events - fatalities (G5 active)



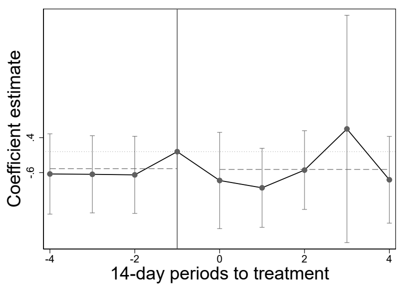
Panel C: Military Operations (no G5) fatalities



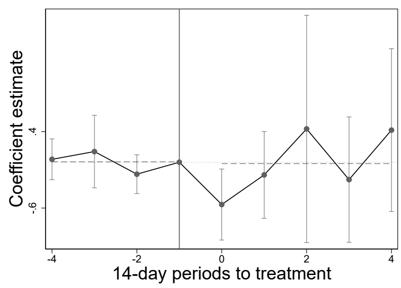
Panel D: Military Operations (G5 active) fatalities



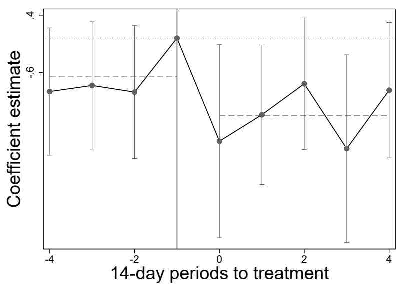
Panel E: Attacks by armed groups (no G5) fatalities



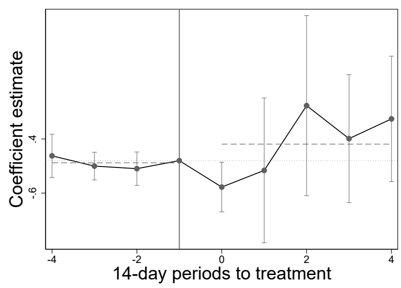
Panel F: Attacks by armed groups (G5 active) fatalities



Panel G: Civilian Casualty fatalities (no G5)



Panel H: Civilian Casualty fatalities (G5 active)



Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2021). The graph show coefficients on time-to-treatment dummies around the trigger operation. The model includes operation fixed effects. Standard errors are clustered at the grid-cell level, and grey bars represent 95% confidence intervals.

Table A7: Reaction to trigger events - by Actor

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Events involving Islamist groups								
Border x Post	-0.10	-0.05	-0.10	-0.05	-0.17	-0.05	-0.16	-0.01
	(0.16)	(0.04)	(0.16)	(0.04)	(0.15)	(0.04)	(0.13)	(0.03)
Border x Post x G5	0.40*	0.01	0.39*	0.08**	-0.01	0.01	0.15	-0.01
	(0.22)	(0.05)	(0.22)	(0.04)	(0.20)	(0.05)	(0.16)	(0.04)
Panel B: Events involving Communal militia								
Border x Post	-0.41	-0.01	-0.09	-0.03	0.15	-0.01	-0.11	-0.02
	(0.39)	(0.04)	(0.10)	(0.03)	(0.32)	(0.04)	(0.08)	(0.03)
Border x Post x G5	0.73	0.01	0.33	0.04	0.01	0.01	0.25	0.01
	(0.48)	(0.04)	(0.22)	(0.03)	(0.37)	(0.04)	(0.20)	(0.03)
Panel C: Events involving Security Forces and civilians								
Border x Post							-0.04	-0.41***
							(0.02)	(0.16)
Border x Post x G5							0.05*	0.62**
							(0.03)	(0.25)
Observations	3168	3168	3168	3168	3168	3168	3168	3168
Clusters	352	352	352	352	352	352	352	352

Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2021). Results are based on the estimating equation presented above. Standard errors are clustered at the grid-cell level and presented in parentheses; stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A8: Reaction to trigger events - alternative criterion

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border x Post	-0.19	-0.09	-0.45	-0.18*	0.27	0.09	-0.28	0.02
	(0.99)	(0.14)	(0.37)	(0.10)	(0.92)	(0.10)	(0.38)	(0.07)
Border x Post x G5	1.08	0.13	0.74	0.19*	0.34	-0.06	0.78	-0.02
	(1.09)	(0.15)	(0.46)	(0.11)	(1.00)	(0.12)	(0.53)	(0.08)
Mean DV	0.836	0.211	0.350	0.071	0.486	0.140	0.344	0.104
Standard Deviation	5.674	0.686	3.964	0.464	3.209	0.420	2.211	0.338
Observations	1827	1827	1827	1827	1827	1827	1827	1827
Clusters	203	203	203	203	203	203	203	203

Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2020) claiming more than 10 lives. Results are based on estimating equation (2). Standard errors are clustered at the grid-cell level and presented in parentheses; stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A9: Reaction to trigger events - alternative ”G5 treatment period”

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border x Post	0.76	-0.10	-0.57	-0.28	1.33	0.18	0.30**	0.06
	(1.52)	(0.23)	(0.65)	(0.18)	(1.40)	(0.16)	(0.14)	(0.09)
Border x Post x G5 (post Sep 2017)	-0.67	0.05	0.79	0.30*	-1.45	-0.24	-0.31	-0.11
	(1.54)	(0.24)	(0.67)	(0.18)	(1.41)	(0.16)	(0.26)	(0.09)
Mean DV	0.693	0.178	0.313	0.054	0.380	0.125	0.360	0.099
Standard Deviation	4.676	0.576	3.373	0.372	2.635	0.387	2.432	0.332
Observations	3168	3168	3168	3168	3168	3168	3168	3168
Clusters	352	352	352	352	352	352	352	352

Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2020) claiming at least 10 lives. Results are based on estimating equation (2). Standard errors are clustered at the grid-cell level and presented in parentheses; stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A10: Reaction to trigger events - allowing for repetition of grid cells

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events	Fatalities (count)	Events
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border x Post	-0.73*	-0.14**	-0.43**	-0.08**	-0.31	-0.06	-0.86***	-0.07*
	(0.38)	(0.06)	(0.18)	(0.03)	(0.34)	(0.05)	(0.26)	(0.04)
Border x Post x G5	1.46***	0.19***	0.59***	0.10***	0.86**	0.09*	1.05***	0.08*
	(0.49)	(0.06)	(0.21)	(0.03)	(0.44)	(0.05)	(0.29)	(0.04)
Mean DV	0.947	0.211	0.404	0.063	0.543	0.148	0.384	0.104
Standard Deviation	6.086	0.668	4.023	0.446	4.103	0.432	2.547	0.347
Observations	7155	7155	7155	7155	7155	7155	7155	7155
Clusters	359	359	359	359	359	359	359	359

Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2020) claiming at least 10 lives. Results are based on estimating equation (2). Standard errors are clustered at the grid-cell level and presented in parentheses; stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A11: Reaction to trigger events - Dummy outcomes

	All events		Military operations		Attacks by armed groups		Attacks on Civilians	
	Fatalities (any)	Events (any)	Fatalities (any)	Events (any)	Fatalities (any)	Events (any)	Fatalities (any)	Events (any)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Border x Post	-0.09**	-0.11**	-0.05**	-0.07**	-0.05	-0.05	-0.07**	-0.07*
	(0.04)	(0.05)	(0.02)	(0.03)	(0.04)	(0.05)	(0.03)	(0.04)
Border x Post x G5	0.07	0.11*	0.06**	0.10***	0.01	0.03	0.06	0.06
	(0.05)	(0.06)	(0.03)	(0.03)	(0.04)	(0.06)	(0.04)	(0.05)
Mean DV	0.097	0.141	0.027	0.037	0.074	0.110	0.063	0.090
Standard Deviation	0.296	0.349	0.161	0.189	0.262	0.313	0.244	0.287
Observations	3168	3168	3168	3168	3168	3168	3168	3168
Clusters	352	352	352	352	352	352	352	352

Notes: Observations at the grid-cell level, binned in two-week periods, in two-month windows around major French operations (2010-2020). Results are based on estimating equation (2). Standard errors are clustered at the grid-cell level and presented in parentheses; stars indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Chapter III

Soldiers versus Laborers: Legacies of Colonial Military Service and Forced Labor in Mali

Joint with Zhexun Mo and Ismaël Yacoubou-Djima

1 Introduction

Coercive labor systems established during colonial rules have been shown to have development impacts persisting until today (Dell, 2010; Dell & Olken, 2020; Lowes & Montero, 2021a; Archibong & Obikili, 2023). The compulsory military conscription system, instituted by the French colonial authorities in Africa, is one of these coercive colonial labor regimes that impacted millions of recruits across French African colonies at the time (M. J. Echenberg, 1991). Its coerciveness was even further amplified during the colonial rule in Mali (formerly known as Soudan Français),[1] where due to local labor shortages, the French colonial authorities resorted to military reservists as forced labor for public infrastructure construction. For an extended period of more than two decades (1927-1950), more than 50,000 reservist forced laborers[2] were sent for public works construction with deplorable working conditions, low salaries and high mortality risks.

The historical functioning of this forced labor regime, as well as its potential long-term legacies on local development still remain largely unknown. This paper investigates both questions using colonial district-level conscription tables and around 180,000 newly digitized individual conscript records from military archives. These records constitute a quasi-exhaustive compilation of military recruitment in the colony of French Sudan over a span of more than two decades. From these soldier files, we extract the localities of origin of the conscripts and the nature of their military drafts, categorizing them as either regular soldiers or reservists.[3] We then match these localities with four rounds of contemporary Malian censuses to identify, for each matched locality, the total share of soldiers and reservists out of all conscripts recruited during the period of time (1927-1950) when military forced labor was in place.

We then exploit this locality-level variation, along with district-level variation in the share of reservists mobilized as forced labor workers, to assess the long-term effects of colonial

[1] In this paper, the terms colonial Mali, French Sudan and Soudan Français are used inter-changeably.

[2] In terms of aggregate number, this is on the same order of magnitude as the total number of regular army recruits over this period of time.

[3] Two other categories, recaptured absentees and volunteers, account for around 5% of the total number of conscripts

exposure to regular military service and military forced labor on contemporary development outcomes. The total share of soldiers of a locality resulted from repeated draws of a lottery which, on an annual basis, randomly selected regular soldiers from a pool of deemed fit young males in each colonial district, in order to fulfil the regular soldier recruitment quota in that given year. Hence, those who were randomly drawn up would serve as soldiers in the colonial army, while the rest remained as reservists. As the average locality sent a total of fewer than 23 conscripts (including both soldiers and reservists) over the full period, the locality total share of soldiers did not necessarily converge to the colonial district share of soldiers, generating some random variation at the locality level within a given colonial district.

This random variation at the locality level thus allows to identify a causal effect of the share of young men enrolled as soldiers versus reservists. However, while soldiers from all districts were assigned to the same military service, the alternatives faced by military reservists varied across districts depending on the share of activated reservists for forced labor work. In districts where almost no reservists were mobilized, the lottery determined whether a conscript became a soldier or remained as a true reservist: avoiding the draw merely meant avoiding military service. In contrast, in districts where all reservists were mobilized for forced labor work, the lottery essentially left individuals with no choice but military service or the prospect of forced labor. Hence, the heterogeneous impact of a locality's total share of soldiers, depending on the district-level activation rate of reservists, allows us to isolate the effect of forced labor, whilst the estimated effect of the total share of soldiers for a null activation rate provides us with estimates of the impact of military service, relative to true reservists.

Our exploration into colonial-era determinants and impacts of forced labor and military recruitment provide useful insights into understanding their long-term effects.

First of all, through population projections, we estimate that active military service and forced labor recruited approximately 6% of young men from cohorts of 20-year-olds between 1927 and 1950. Despite constituting a significant proportion of the population during the colonial era, the relative demographic weight of these exploitative systems suggests that any persistent effect likely occurred through informational and normative channels rather than solely demographic shocks.

Secondly, historical analysis reveals a positive correlation between the activation of reservists for forced labor and the patterns of voluntary enlistment into the army. This suggests that forced labor was markedly perceived as an adverse experience during the colonial period, prompting young men to actively evade it by opting for enlistment in the regular army.

Finally, our investigation of the determinants of army recruitment at the colonial district and locality levels suggest that the colonial authority acted as a rational actor optimizing its levels of coercion and organization of labor, and did so in a centralized manner. This confirms the validity of our empirical strategy based on lottery rates. At the colonial district level, the intensity of military conscription and the allocation of conscripts into soldiers, true reservists and forced laborers were closely correlated with district-level labor demands and ease of recruitment, indicating a very centralized recruitment procedure. Conversely, at the locality level, the within-district distribution of conscripts between soldiers and reservists appeared to be independent of locality-level characteristics and was not auto-correlated over time or space, aligning with historical evidence that it was the result of a random allocation across localities within colonial districts.

Turning to the long-term consequences of military recruitment, our findings suggest an average null effect of the total share of soldiers on a locality's contemporary development outcomes. However, this masks highly heterogeneous effects among villages (localities) located within former colonial districts with high and low average forced-labor activation rates of reservists. This heterogeneity suggests that against a "true reservist" counterfactual, both localities affected by recruitment in the regular army or in the forced labor camps were disadvantaged over the long run. These effects are particularly strong in terms of human capital accumulation, measured by the share of village residents having attained primary schooling. They are also consistently significantly across four waves of post-colonial Malian censuses over a timespan of more than 30 years, indicating a possible mechanism of intergenerational transmission of educational and economic disadvantages, although the effects on school enrollment dissipates over time and is no longer significant in 2009.

We hypothesize that exposure to colonial military forced labor could have led to increased rejection of colonial institutions or those resembling colonial setups among indigenous populations, including the modern schooling system. Our findings using Afrobaromter data collected in 2018 indicate that localities affected by colonial conscription were more likely to reject the French military intervention in Mali. While this observation supports the hypothesis that colonial coercion generated resentment towards France, further investigation is necessary to ascertain whether this sentiment extended to another colonial legacy: formal schooling. The diminishing effect observed after the 1991 democratization in Mali and subsequent education reforms, which granted greater autonomy to schools, lends credence to this hypothesis. However, additional research is warranted to fully explore this relationship.

Our paper contributes to different strands of literature in the following ways.

First of all, the literature on understanding the historical legacies of colonial forced labor has regained momentum in recent years in economics with the possibility of digitizing massive historical datasets. On the one hand, Nunn and Wantchekon (2011) show that slave trade had persistent impacts on economic development through lower interpersonal trust among ethnic groups most exposed to the slave trade in Africa, whereas Dell (2010) finds that more formal institutional channels such as land tenure and public goods provisions explain the negative developmental effects of forced mining labor system in Peru under the Spanish Empire. More recently, forced labor was uncovered to be a key instrument in financing both the French Colonial Empire in Africa (Van Waijenburg, 2018), as well as the British colonial government in Nigeria (Archibong & Obikili, 2023). In the latter case, the use of prison labor even resulted in lower trust in contemporary judicial systems. Ultimately, in former Belgian African colonies, Lowes and Montero

(2021a) show that forced-labor concession systems in colonial Belgian Congo led local chiefs to be less accountable and provide fewer public goods, yet they also made indigenous locals more trusting of each another due to such collective trauma, and Blouin (2021) shows that ethnicity-targeted colonial forced labor practices might have resulted in lower inter-ethnic trust and collaboration in Rwanda and Burundi.

Despite this aforementioned burgeoning literature, developmental repercussions of colonial military conscription and the forced labor entailed have been under-studied, notwithstanding their historical and contemporary significance (M. J. Echenberg, 2009). Colonial-time Mali was the only French (African) colony where military reservists were systematically activated as forced labor for an extended period of more than two decades from 1927 until 1950 (M. Echenberg & Filipovich, 1986). Given that these individuals originally had the possibility of having military careers, the unexpected activation of such forced labor could have produced very different long-term legacies compared to more conventional forced labor systems. To the best of our knowledge, we will be the first study to assess empirically the long-term developmental legacies of colonial military forced labor at a fine-grained village level in Africa. Although there has been an extensive literature in history documenting the widespread practice of various forms of coercive labor, as well as their significance and necessity for the colonial economy in former French Africa (Fall, 1993; Bogosian, 2002, 2015; Van Waijenburg, 2018; Tiquet, 2019, 2020; Cogneau, 2023), there is rather little or scant evidence on the long-term legacies of such forced labor regimes on contemporary economic development in former French African colonies, with the exception of Dupas et al. (2023), where the authors analyze the impact of forced labor on internal migration and traditional family structures in colonial Burkina Faso (ex Upper-Volta). Our study further contributes to this strand of literature by assessing the long-term developmental repercussions of another equally important forced labor regime at the height of French colonial conscription in Mali.

Within the broader literature on the persistent effects of exposure to colonial coercion, studies have highlighted lasting legacies on attitudes towards modern institutions such as healthcare or justice systems (Lowes & Montero, 2021b; Archibong & Obikili, 2023). Our paper contributes to this literature by documenting how negative experiences with one colonial institution, the army, can spill over to another institution, the schooling system, particularly in a setting where both institutions trace their origins back to the colonial period.

Our project also relates to the themes of economic inequalities and social mobility, as the structures of military forced labor affected localities of Mali unequally, which might have reinforced (or alleviated) pre-existing within-village inequalities.[4] As such, the specific nature of such extractive colonial policies might serve as another potential determinant of intergenerational mobility in Sub-Saharan Africa today (Alesina et al., 2020, 2021).

Ultimately, our paper also speaks to the growing literature investigating the long-term impacts of conscription and military service (Vanden Eynde, 2016; Libois et al., 2020; Ronconi & Ramos-Toro, 2022). The papers the most closely related to ours are Salem (2022) and Salem and Seck (2022), where the authors also looked at the historical event of French colonial conscription in Africa. However, their focus differs from ours in two important dimensions: first of all, they focus on French colonial military recruitment in North Africa, which was a volunteer-based system, while in French West Africa military conscription was compulsory; secondly, they look at the impact of overseas migration associated with military out-posting on local development, while we focus on the differential impacts of being assigned to different segments of the colonial army, with a specific focus on the potential negative ramifications associated with reservist forced labor.

The remainder of this paper is structured as follows. Section 2 provides the historical background of military conscription and the forced labor entailed in colonial Mali and French West Africa. Section 3 discusses our main data sources as well as summary statistics. Section 4 provides historical empirical analysis of the forced labor system, and Section 5 presents main results on long-term impacts on contemporary development outcomes. The final Section concludes.

[4]Some anecdotal historical evidence indicates that disadvantaged groups (casted people and former captives) were potentially more likely to be sent for forced labor, and more advantaged individuals, such as chiefs' sons, would be more likely to be formally selected by the drafting commission into the regular army (Mann, 2006).

2 Historical Background

French colonial authorities had been recruiting indigenous African soldiers ever since the onset of their conquest in West Africa in the 19^{th} century.[5]

The first great mobilization took place during the First World War, when over 170,000 soldiers were levied across the entire French West Africa and over 30,000 of them died for France (M. J. Echenberg, 1991). After the war, having witnessed the potential to mobilise indigenous young males for labor needs, the French colonial authorities decided to institutionalize military recruitment via a compulsory peacetime conscription system starting circa 1919, where every year a fixed number of soldiers would be recruited across the entire federation.[6]

Within colonial Mali, when the colony head received the annual military quota from the General Governor of French West Africa, he would distribute it across the colonial districts[7] within Mali with some regards to local population densities (Cogneau & Mo, 2023), then the district administrator would be in charge of fulfilling this military quota via implementing a mobile drafting commission board within the district in that given year.[8]

While fulfilling the military target, the colonial authorities came to the realization that the number of local young males deemed fit by the mobile drafting board in a given

[5]The early recruits often involved the purchase of groups of slaves and war captives (Zuccarelli, 1962; Renault, 1972; Klein, 2011), and were mainly for the purpose of either conquering further inland or guarding newly-obtained territories. Such African troops started to be referred to as **Tirailleurs Sénégalais** with the first permanent unit of black Senegalese soldiers recruited in 1857, and the term had been used to refer to all colonial French soldiers composed of indigenous black African recruits, either they were conscripted from French West Africa (Afrique Occidentale Française) or French Equatorial Africa (Afrique Equatoriale Française).

[6]Military conscription in the colonial federation of French West Africa was a very hierarchical system, where the military quota was decided annually in Paris and then sent for execution at various levels of administration in the French African colonies. For an illustration of this top-down procedure please see Figure A1.

[7]A colonial district ("cercle" in French) is the administrative layer lower than that of a colony. It is sually referred to as a colonial district, headed by a district administrator (Delavignette, 1939; Huillery, 2009; Do et al., 2020). In colonial Mali, there were around 20-22 colonial districts in total during our time period of study (1927-1950), with the fluctuations due to some merging and splitting of districts over time.

[8]Besides the district administrator, main members of the mobile drafting board also include some French army officers and military physicians. The board would travel from one district to another to fulfill the total recruitment quota within a given colony (in our scenario Mali).

district often far exceeded the military quota assigned to that district. As such, the colonial administration drew via a random lottery a relatively small fraction of these fit young males to fulfil the recruitment target, and the rest would then be classified as military reservists.[9]

During the interwar peacetime period, these reservists did not serve much purpose for the army. However, given endemic labor scarcity in colonial Africa (Fenske, 2013), as well as the initiation and renovation of large-scale infrastructure projects across colonial Mali in the 1920s, the needs for readily available fit labor increased. Consequently, the hitherto idle military reservists were viewed as an untapped source of cheap labor and started to be called into service for such public infrastructure projects. Formally, a decree was issued by the Governor-General of French West Africa to bring the reservists into (forced) labor service on December 4^{th}, 1926. Subsequently, in 1927 a first batch of 1,000 young indigenous males within the reservist camp were activated as forced labor in colonial Mali, with half of them sent for the newly created *Office du Niger* irrigation project, and the other half for renovating the *Kayes-Bamako* railways.[10]

An illustration of the whole activation procedure of forced labor within the reservist camp is provided in Figure A2a and Figure A2b, where it could be seen that in a given district for a given year, after the medical examinations were performed among the eligible young males at the conscription center, and that the volunteers to join the army were subtracted from the deemed fit young males,[11] a random sub-share of the fit men would be categorized (via a lottery) as first-portion regular army recruits to match the military quota

[9]In colonial military recruitment terms, the regular army recruits were referred to as "first-portion" soldiers and the reservists as the "second-portion" in the military contingent, who would be mobilized for regular army service on special occasions such as war times.

[10]*Office du Niger* is a large-scale irrigation project within the Niger River Delta in the current-day Segou region of Mali. The forced laborers were called upon mainly to excavate canals and build the gigantic dam located in *Markala*, Segou. *Kayes-Bamako* railway is the segment of the *Dakar-Bamako* that lies within colonial Mali. Together with *Office du Niger*, they became the two most important destinations of reservist forced labor.

[11]The number of volunteers was capped in the peacetime conscription system in French West Africa. As it's written by M. J. Echenberg (1991): *"The French gradually imposed a limit on the number of volunteers that could be permitted from any one recruitment district. This ceiling came to be one-third of the contingent; as a result, in some areas volunteers were actually turned away, whereas elsewhere a significant number of reluctant soldiers were being conscripted."*

assigned to that district in a given year, and the rest would be classified as second-portion reservists.

Specifically, this random allocation of fit young males into either one of the segments of the military army is detailed below by M. J. Echenberg (1991): *"After the volunteers were subtracted from the quota a lottery was held to determine the balance. Those who drew a "bad number" went into the army as conscripts; they were called the first portion. The remainder whose numbers had not been drawn became part of the second portion."*[12] However, for the military reservists (second portion), their lucky draw did not appear to be so lucky during the years of forced labor. Specifically, given the labor needs in a given year, a fraction of them would be activated and respectively sent for either *Office du Niger* or the railways.[13]

On average, around 2,000 to 3,000 reservist forced laborers were exploited on an annual basis. They were organised to work and live in camps on the public works sites for 2 to 3 consecutive years.[14] Such a military forced-labor system was in place for more than two decades across the colony of Mali, up until the end of the 1940s when the abolition of forced labor was gradually implemented across all French African colonies (Ginio, 2017). Until then, more than 50,000 reservists had been subject to forced labor over this extended period of time.

The mistreatment of these reservist forced laborers in Mali has been well-documented in the vast literature in history (Magasa, 1978; M. Echenberg & Filipovich, 1986; M. J. Echenberg, 1991; Fall, 1993), and has left an important imprint in the collective memory of the Malian society (Bogosian, 2002, 2015). It was a system associated with not only high mortality rates, but also the stigma of serving as forced labor while being labelled as

[12] The lottery rate is not an exogenous measure itself, as it has been shown in Cogneau and Mo (2023) that it would be adjusted in response to the military quota assigned to a given district in a given year. However, after the rate was determined, which fraction of young males would go into the second portion, and the rest into the first portion was purported to be quasi-random as described above. This quasi-random draw would be exploited later in our empirical strategy for long-term persistence analyses.

[13] The activation rate of forced labor among the reservists differ from one year to another, which gives us variation at the district-year level over time in colonial Mali.

[14] After the termination of this two-to-three-year contract, the forced laborers usually would return back to their places of origin. M. Echenberg and Filipovich (1986) suggest this was the case for more than 80% of them.

a "formal" segment of the military.[15] Furthermore, such traumatic experience of forced labor was even more contrasted with the experience of the young males classified as first-portion regular soldiers, who enjoyed higher compensation, more honorable status as well as pension payment after retirement of service, all of which were non-existent for the reservist forced labor workers.[16]

On the other hand, the circumstances and treatments of regular-army African soldiers underwent distinct shifts throughout our period of study. Before the onset of the Second World War, the colonial army was in charge of policing metropolitan France and its colonial offshoots in North Africa and the Levant. During the war years, the mortality rates among African soldiers surged to an estimated 12% (M. Echenberg & Filipovich, 1986). In the post-war period, repatriation and compensation delays for those surviving prisoners sparked the Thiaroye revolts, which was violently suppressed and claimed the lives of 35 soldiers (Mann, 2006). Soldiers enlisted during the Vichy era were deployed in Mali to deter potential British incursions into the Germany-allied Vichy-controlled Sudan. In 1943 and 1944, they were mobilized within the Free French Forces under Marshal Lattre de Tassigny's command. Those African soldiers were however deprived of the spoils of victory as De Gaulle issued an order for their demobilization before the liberation of Paris. This "whitening" of the French Army, revealed the lack of recognition of the important role played by African soldiers in France's victory in World War Two.

In this paper, we intend to first empirically test the historical significance, as well as the internal functioning of such a military forced labor system in colonial Mali; and second of

[15]For instance, such mistreatment was documented by M. J. Echenberg (1991) during his interview with one of the former reservist forced laborer in Mali: *"To put it bluntly, it was labor at the cheapest possible price. We were paid 30 francs CFA a month. We were poorly fed, but we received work clothes as if we were really soldiers ... Any Malian or Voltaique family still has a fresh memory of the building of the bridge (at Markala) where men worked in the rain and under the whip without respite. The lazy and the revolutionaries were thrown into the river to intimidate the other workers...*

[16]For instance, Bogosian (2002, p. 18) explained in detail the different career and life trajectories faced by the reservists (deuxième portion) and regular soldiers (première portion): *"At the moment of recruitment, the young men who were chosen for the deuxième portion and for the army were no different from each other; once each group began their service, their experiences were enormously different. While soldiers learned to handle rifles, men in the deuxième portion took up picks and shovels."* The potential reasons for such differences are also noted in another related study: *"As more and more men joined the deuxième portion service during the war years, they became ever more aware of the sharp differences between their work conditions and lack of benefits when compared with the rewards for the more honorable and prestigious work of soldiers."* (Bogosian, 2015)

all, we plan to gauge if such colonial-time exposure to the draconian forced labor regime (as well as staying as a true reservist) versus the regular army service have led to differential development trajectories across Malian villages until this day.

3 Data

3.1 Conscription and Forced Labor Variables

3.1.1 CAPM data : individual-level conscript files

Our conscription data at the locality level were obtained by digitizing hand-written individual-level soldier files stored at the French Archive Center of Military Personnel (CAPM) in Pau, France. An example of an anonymized individual soldier file is provided in Figure B1.

Each file contains detailed information about the individual conscript, such as his first and last names, years of birth, villages and districts of birth, as well as those of residence, parents' names and their village of residence, nature of draft, year of draft, height, facial features, profession, marital status, and more. For this paper, the variables of particular interest are related to the villages and districts of birth and residence, the nature and year of the draft. From June 2021 to February 2023, pictures of the entire sample of individual files were taken at the CAPM for the time period of 1927 till 1950.[17]

The subsequent step in the digitization process consisted in the automatic transcription of the different fields of interest on the individual files. Rather than following a manual data entry process, we opted for the automatic transcription of the files.[18] Details of the transcription process are given in the appendices Section A.4.

In total, after data cleaning, we obtained 178,430 individual soldier files. The breakdown

[17]We shared some of the costs of photo-taking with another team of researchers who were interested in some of the files we were collecting and were also engaging in data collection at CAPM at the same time.

[18]Opting for an automatic transcription of the files has proved to be cost-efficient and helped us comply with privacy requirements of the CAPM administration. We are also grateful for a large portion of the training data that came from data collection effort by a team led by Denis Cogneau and Alexander Moradi.

of the files by military draft category is given in Table 1. We classify almost all (except for less than 0.50 percent) of the files in terms of the military draft category. Among the successfully classified files, we find that approximately 29 percent of our sample consisted of soldiers drafted into the regular French army. The majority of the sample, around 69.9% percent of the individuals, were classified as military reservists. We are able to distinguish "true" reservists from reservists activated as forced laborers for only two years (1941 and 1942) at the individual level. Figure 1 provides a breakdown of the different military draft categories by year while Table 1 summarizes the distribution for the full period. The bottom panel of Table 1 indicates that for 1941 and 1942, around 70 percent of reservists were activated as forced labor, which is higher than the average activation rate observed in the district-level data but consistent with historical evidence that use of forced labour was especially strong during the years of the Vichy regime in France (Tiquet, 2019).

We use the handwritten information on the village of residence and birth transcribed from individual conscript files to match them with the names of localities[19] listed in the 1976 general population census of Mali.[20] Due to the potential spelling discrepancies introduced by the colonial administration agents and the machine-learning-based text recognition process, we employ a fuzzy matching procedure. The details of this matching procedure are elaborated in the appendices, Section A.5.

Table 2 provides a comparison of the individual sample that is matched versus the entire individual sample. It shows that our sample overestimate slightly the number of reservists. However the difference in shares in comparison to the entire sample remains under one percent, which we do not think will considerably bias our results especially if it happens in a similar way across all localities that we are able to match. Table C4 in Appendix further compares the matching rate across colonial districts. Individuals from Northern districts the country (Bourem, Gao, Goundam, Gourma and Tombouctou) are one average much less well matched than the average (column 1) and this creates a

[19]Within the context of Malian censuses, a 'locality' is synonymous with a 'village.' Hence, the two terms will be used interchangeably throughout this document.

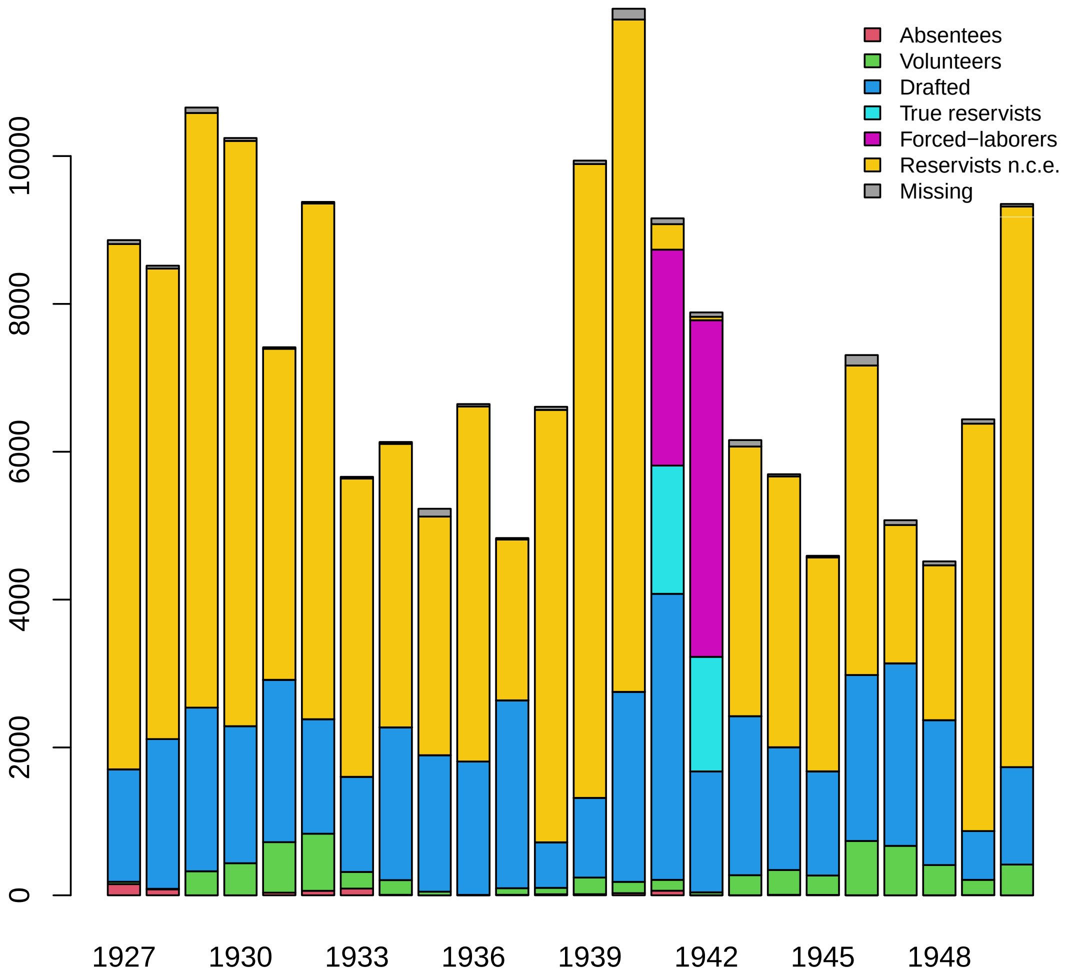
[20]This census provides the earliest and most comprehensive list of localities available to us. We are in the process of exploring earlier village lists to better understand the evolution of village creation in Mali over time, as well as to evaluate the accuracy of the matches obtained to date.

Table 1: Summary count of military draft categories in the CAPM data

Military draft category	N	%
Entire Sample		
Soldiers		
Absentees	569	0.32
Volunteer soldiers	6,687	3.8
Drafted soldiers	45,466	25.5
Total soldiers	52,722	**29.6**
Reservists		
Total Reservists	124,671	**69.9**
Missing or unclassifiable	850	**0.5**
Total	178,243	**100.0**
1941 - 1942		
Soldiers		
Absentees	61	0.4
Volunteer soldiers	165	1.0
Drafted soldiers	5,686	33.5
Total soldiers	5,912	**34.9**
Reservists		
True reservists	3,278	19.3
Forced-laborers	7,475	44.1
Reservists n.c.e.	207	1.2
Total Reservists	10,960	**64.6**
Missing or unclassifiable	88	**0.5**
Total	16,960	**100.0**

Notes: An observation represents a conscript file. The first panel is the entire sample of all files. The second panel shows the breakdown of reservists between reservists activated as forced laborers and true reservists which did not serve. We are able to provide this breakdown for the years 1941-1942. For some files (categorized as Reservists n.c.e), it was not possible to classify in the above two categories

Figure 1: Count of the number of files by category of military draft by year in the CAPM data



Notes: We are only able to classify the true reservists from the forced laborers in 1941 and 1942. Pictured is the count of the full sample. For a breakdown between the matched and unmatched sample with the censuses data, please see Table 2

Table 2: Military draft categories matched and unmatched sample

Military draft category	Matched N	Matched %	Entire Sample N	Entire Sample %	Matched-Entire Sample Difference (%)
			Entire sample		
Soldiers					
Absentees	435	0.31	569	0.32	−0.007
Volunteer soldiers	5,091	3.65	6,687	3.75	−0.098
Drafted soldiers	34,715	24.92	45,466	25.51	−0.591***
Total soldiers	39,806	28.57	52,153	29.26	−0.689***
Reservists					
Total reservists	98,479	70.68	124,671	69.94	0.738***
Missing or unclassifiable	606	0.43	850	0.48	−0.042*
Total	139,326	100.00	178,243	100.00	
			1941 - 1942		
Soldiers					
Absentees	44	0.33	61	0.36	−0.033
Volunteer soldiers	130	0.97	165	0.97	−0.007
Drafted soldiers	4,423	32.85	5,686	33.53	−0.680
Total soldiers	4,553	33.81	5,851	34.50	−0.688
Reservists					
Forced-laborers	6,028	44.76	7,475	44.07	0.690
True reservists	2,639	19.60	3,278	19.33	0.270
Reservists n.c.e	135	1.00	207	1.22	−0.218*
Total Reservists	8,802	65.36	10,960	64.62	0.742
Missing or unclassifiable	67	0.50	88	0.52	−0.021
Total	13,466	100.00	16,960	100.00	

Notes: An observation represents a conscript file. We do t-test mean score of equality of the mean across the two samples. *** p<0.01, ** p<0.05, * p<0.1

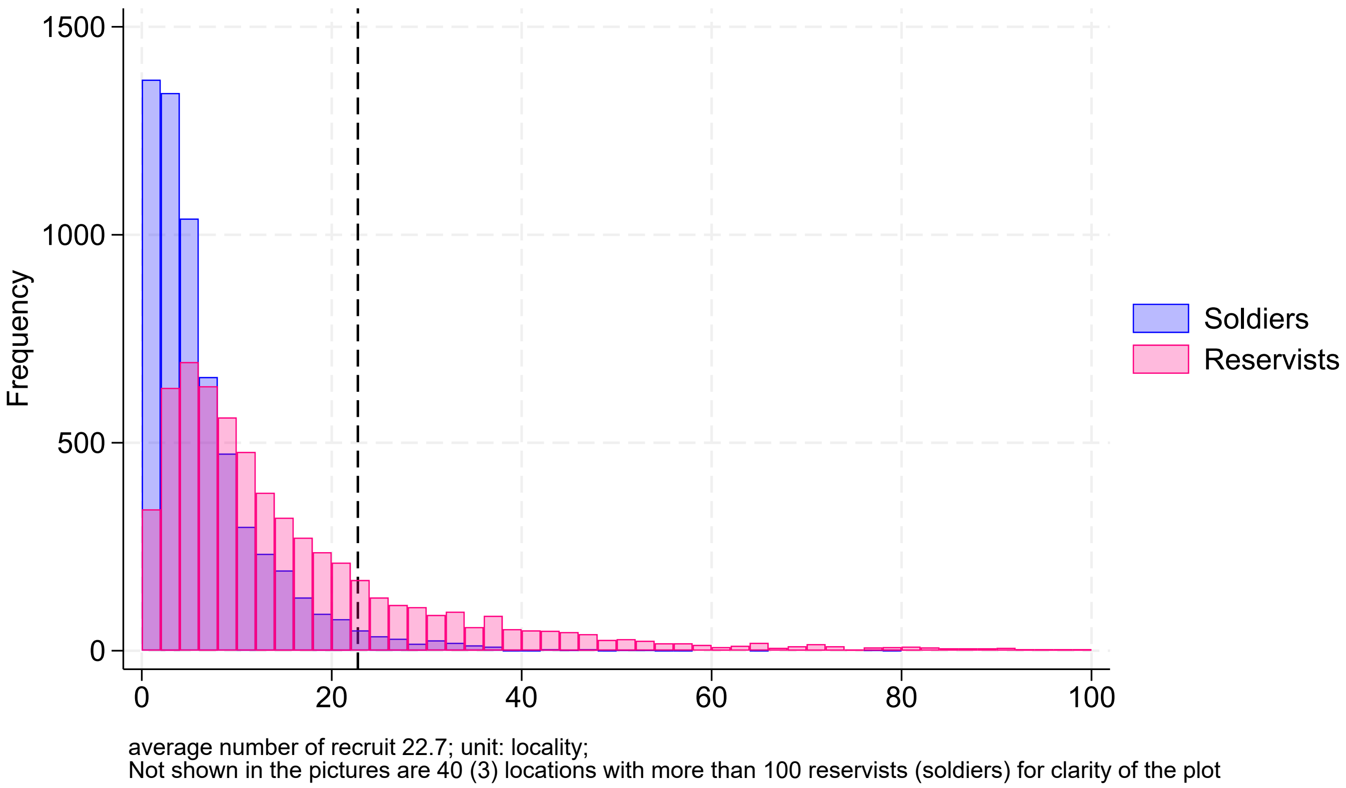
positive correlation between the share of reservists activated as forced laborers and the matching rate (column 2), as reservists from the Northern districts were much less extensively activated for forced labour. Excluding these districts from the sample results in a null correlation between district level activation rate and matching rate (column 3).

Thanks to the fuzzy matching results, we are able to examine the spatial distribution of the share of reservists using locality coordinates collected as part of the work done in Mesplé-Somps et al. (2018). This could be observed in Figure B3, where we plot the cross-village variations in the share of soldiers within each colonial district in Mali, and we do observe a quite wide-spread distribution of the locality-level share of soldiers across the board, ranging from as low as 10 to 20% to sometimes reaching 70 or 80% in some localities. We also see on Figure 2 that many localities do not have any soldiers enlisted during the period of interest, as the average locality sent less than 23 conscripts. Such small numbers of total recruitment at the locality level also gives us enough variation, when the district-level lottery rate (share of soldiers) is calculated at the locality-level within a given district, due to sampling errors in small samples.

3.1.2 ANS data: district-level conscript tables

In addition to the individual-level data, we also obtained colonial district-level conscription data digitized in Cogneau and Mo (2023). These data were taken from the colonial district-level annual military reports in Section Série-4D at the National Archives of Senegal (ANS) in Dakar. The military reports contain conscription tables with detailed information on the recruitment procedures of the mobile drafting boards. Specifically, as Figure B2 shows, they include information on the recruitment quota assigned to a given district, the total number of enumerated young males, absentees, deemed fit individuals, regularly drafted soldiers, volunteers & reservists, etc. Summary statistics of the ANS district-level data is provided in Table B1. In an average-sized district in colonial Mali, the overall recruitment of military reservist is around 270 young males, as opposed to only 110 males recruited into the regular army. The average volunteering rate is around 15%, which was also capped as discussed previously that the French colonial authorities

Figure 2: Histograms of the count of reservists and soldiers over the entire period. *Notes:* Not shown in the pictures are 40 (3) locations with more than 100 reservists (soldiers) for clarity of the plot



didn't wish to have an army completely composed of volunteers. We further observe that the deemed fit rate was only around 20% over time, which means that the overwhelming majority of enlisted and present young males were declared unfit for the military. This also means that the individuals of interest in our paper, either regular army recruits or reservists, were deemed fit individuals by colonial medical examination standards, as such any long-term impacts we may uncover later on would be driven by the exposure to forced labour and military service of a selected group of the most able-bodied young males in their respective villages in colonial Mali.

Despite having collected the data from two independent sources across time and space, we found that there is a very close correlation between the two datasets in terms of the total number of military conscripts in the entire colony of Mali across years as depicted in Figure C1 in Apendix in section A.7.[21]

3.2 Outcome Variables - Descriptive Statistics

We use the 1976, 1987, 1998 and 2009 Malian censuses[22] to obtain measures of economic development at the locality level. We calculate locality-level measures of educational attainment (shares of individuals 6 years or older, without schooling, with primary schooling, with secondary schooling, with university-level schooling, etc), wealth indexes,[23] within-locality level of inequality.[24] We define fertility as the average number of children in the age range of 0 to 5 out of total number of women in the age of range of 15 to 50.

In addition, for all four rounds of censuses, we also have counts of different types of infrastructures (such as the number of schools, health facilities, dispensaries, etc) at the locality level which could be used as an indicator for local development.[25] Tables B7, B8, B9, B10 provide summary statistics of the different variables we obtained from the cen-

[21]This close correlation applies to the entirety of the CAPM database, not just the individuals matched with the 1976 census.

[22]We thank the National Institute of Statistics in Mali for sharing with us the censuses and allowing the team of researchers to share the data with us for this project.

[23]They are constructed based on a principal component analysis of housing characteristics.

[24]It is measured by the absolute Gini coefficient based on the household-level wealth index calculated previously.

[25]These infrastructures data were previously utilised in Chauvet et al. (2015) and Bernard et al. (2017).

suses. The 1976 individual census appears to have more missing data at the individual level but summary demographic statistics are available at the locality level, and we only rely on those in our analysis. We rely on the matching of localities across censuses conducted by Mesplé-Somps et al. (2018), and we link those localities further back to the ones in the 1976 census. There is little attrition across censuses and we keep for our analysis only the villages that are concurrently present in all four rounds of the censuses. This represents a total of 7,418 villages. Out of these villages, we are able to match 5,613 of them with the CAPM database. Table 3 shows that the villages we are able to match are different demographically from the rest of the villages in the censuses. They have a larger population and have slightly better human capital measures and average wealth values. The demographic bias exhibited in Table 3 can be rationalized, as we expect to match villages that have continued existing throughout colonial times until now, thus they could have had the chance to grow more than newly created villages. On the other hand, as can be seen on Figure 3, it appears that there is no significant geographical bias in terms of the villages we are able to match: the villages in our final sample are rather spatially distributed in an evenly manner across the entire country of Mali.

To further investigate the impacts of colonial military conscription on education, we rely on a school census from the Malian Ministry of Education from 2021 that lists a total of 13619 private, public, Quranic and community schools schools in Mali in about 5000 different localities that we match with the census and archives localities. We can match up 91% of community schools and 87% of public schools, but only 58% and 49% of Quranic schools and non-religious private schools.

We also pool the 7 rounds of geo-coded Afrobarometer data collected in Mali between 2001 and 2018, and match each of the geo-coded enumeration area to all localities within a 5km radius of the coordinates. We use the variable "trust in traditional leaders" collected in rounds 2, 4, 6 and 7 and a variable on the perception of the Barkhane intervention, only available in round 7 [26], to measure the long-term effect of military conscription on perception of customary leaders and the former colonial power

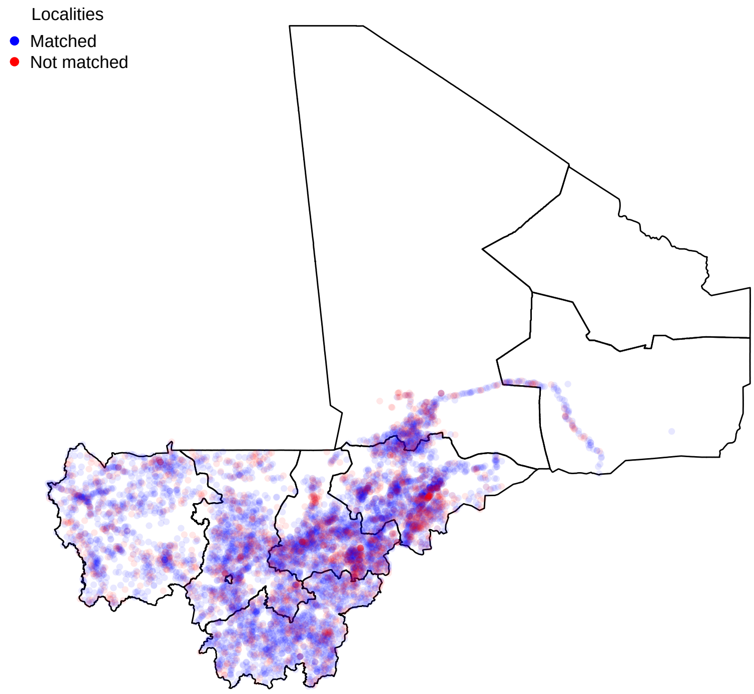
[26] Rounds 8 and 9 should be integrated in our analysis in the future.

Table 3: Outcomes of censuses variables in matched and unmatched localities

	All	Matched	Unmatched	Difference
		Census 1976		
Number of dispensaries	0.077	0.090	0.034	0.057***
Number of schools	0.103	0.115	0.065	0.050***
Number of markets	0.133	0.147	0.090	0.057***
Number of pharmacies	0.022	0.027	0.006	0.021***
Number of Hhs	122.967	130.635	99.125	31.510***
Total residents	621.185	661.656	495.354	166.302***
Female residents	319.383	340.460	253.849	86.611***
Male residents	301.802	321.196	241.505	79.691***
N	7418	5613	1805	
		Census 1987		
Number of Hhs	130.439	138.840	104.314	34.526***
Total residents	689.207	733.936	550.115	183.821***
Female residents	359.554	383.148	286.183	96.965***
Male residents	329.653	350.788	263.931	86.856***
No education	0.941	0.940	0.944	−0.004**
No formal education	0.944	0.943	0.947	−0.004**
Primary school	0.054	0.055	0.051	0.004**
Secondary school	0.002	0.002	0.002	0.0001
University	0.000	0.000	0.000	−0.00001
Average welfare index	-0.176	-0.153	-0.250	0.097***
Absolute within locality gini	0.165	0.165	0.165	−0.001
Inter-quartile range	0.344	0.340	0.356	−0.015
N	7418	5613	1805	
		Census 2009		
Number of Hhs	206.663	219.557	166.566	52.991***
Total residents	1269.345	1351.659	1013.376	338.283***
Female residents	642.706	685.462	509.745	175.718***
Male residents	626.640	666.196	503.631	162.565***
No education	0.770	0.767	0.780	−0.013***
No formal education	0.777	0.773	0.788	−0.015***
Primary school	0.208	0.212	0.197	0.015***
Secondary school	0.010	0.011	0.010	0.0001
University	0.004	0.004	0.005	−0.001**
Average welfare index	-0.419	-0.410	-0.445	0.035***
Absolute within locality gini	0.140	0.142	0.132	0.010***
Inter-quartile range	0.284	0.288	0.270	0.018**
N	7418	5613	1805	

Notes: An observation represents a locality. We compare using a T-test the mean of locality level characteristics in the panel of localities across 1976-2009 census that is matched to the CAPM data and the one that is not matched. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Figure 3: Matched and not matched localities in the 1976 census *Notes:* We are only able to plot localities with coordinates in the census (70 percent of them).



4 Historical Analyses of Conscription and Forced Labor

Was the conscription system effective in mobilizing the local population to meet the military and construction needs of colonial authorities? Analysis of historical data suggests that the colonial authority acted as a central rational actor, selecting the healthiest individuals while limiting draft evasion, and optimizing its levels of coercion and organization of labor in a centralized manner based on its understanding of local characteristics. Analysis of the determinants of conscription outcomes together with analysis of short-term reactions to forced labour suggest that the local population responded to the conscription depending on the level of coerciveness (activation rate) and on its ability and incentives to evade conscription.

4.1 Definition of Key Variables of Interest

Before beginning our analysis, it is helpful to recall the different military draft categories and define our key variables of interest. As discussed in Section 2 and illustrated in Figure A2, the military recruitment procedure started with asking the enumerated eligible young males to present themselves at the conscription centers, and then a medical examination would be performed on the present young males, to decide whether they were fit for military service or not.[27]

In accordance with the procedure of military recruitment discussed in Section 2 and as illustrated in Figure A2, the military recruitment procedure commenced with the summoning of eligible young males to report to the conscription centers, which effectively makes it plausible to calculate the present rate $PreRate$ (share of present young males out of all enumerated young males). Subsequently, a medical examination was conducted to determine their fitness for military service, with a share $FitRate$ (share of deemed fit young males out of all present young males) considered as fit. In theory, after the medical examination, all the young males who were deemed fit were recruited into military ser-

[27]To these individuals, some absentees recaptured from the previous years of conscription would be added. In practice, these absentees represent only a few hundred individuals per year so that we don't focus on these in our analysis.

vice, but for different purposes depending on the units they were assigned into. Overall, we use the term $Conscripts$ to refer to all the deemed fit individuals recruited in military service.

$$\text{Conscripts} \approx \text{Enumerated} * \text{PreRate} * \text{FitRate}^{28}$$

$Conscripts$ could be recruited either as $Soldiers$ or $Reservists$. $Soldiers$ could either be $Volunteers$, who in principle volunteered to join the regular army before the lottery draw, or $DraftedSoldiers$,[29] who were randomly drawn into the army after the volunteering procedure, in order to fulfill the military quota set at the colonial district level for a given year by the Lieutenant Governor of Mali, such that the total number of $Soldiers$ should converge to the military prescribed for that colonial district in that given year. $Conscripts$ who did not volunteer to join the regular army and were not randomly drawn to be $DraftedSoldiers$ either, remained as $Reservists$. Those reservists could either be $TrueReservist$ or $ForcedLaborers$, with the latter ones either serving for constructing irrigation dams at the *Office du Niger*, called $Irrigation$ forced laborers, or for maintenance work on the *Kayes-Bamako* railways, called $Railways$ forced laborers.

These definitions can be summarized as such :

$$\begin{aligned}
&\text{Conscripts} = \text{Soldiers} + \text{Reservists}\\
&\text{Soldiers} = \text{Volunteers} + \text{Drafted Soldiers (Regular Soldiers)}\\
&\text{Reservists} = \text{True Reservists} + \text{Forced Laborers}\\
&\text{Forced Laborers} = \text{Irrigation} + \text{Railways}
\end{aligned}$$

[28] $Enumerated_{ict}$ refers to the total number of enumerated individuals; $PreRate_{ict}$ refers to the present rate (present males out of all enumerated individuals); $FitRate_{ict}$ refers to the Fitness Rate (judged fit people out of all present individuals).

[29] The term "drafted soldier" and "regular soldier" would be used interchangeably in this paper, as they both refer to non-volunteering regular soldiers recruited via the lottery draw.

Our analysis will mostly focus on two steps of this recruitment process: the lottery rate distributing conscripts between drafted soldiers and reservists, and the activation rate, distributing reservists between true reservists and forced workers.

The lottery rate, or soldier share variable can be constructed for each colonial district and each locality on an yearly basis (as long as the total number of military conscripts from a given locality or district was non-zero).

$$\text{Share\_Soldiers}_{iy} = \frac{\text{Drafted Soldiers}_{iy}}{(\text{Reservists}_{iy} + \text{Drafted Soldiers}_{iy})} \tag{1}$$

for year y, in a given village (locality) i, or a given district d.[30]

It could easily be seen that the share of reservists is simply the complement of the share of soldiers in one, where:

$$\text{Share\_Reservists}_{iy} = \frac{\text{Reservists}_{iy}}{(\text{Reservists}_{iy} + \text{Drafted Soldiers}_{iy})} = 1 - \text{Share\_Soldiers}_{iy} \tag{2}$$

Ultimately, given that forced labor workers were in reality *activated* among the military reservists, we also construct the forced-labor activation rate as the number of forced labor workers out of total reservists at both locality and colonial district levels.

$$\text{Activation}_{iy} = \frac{\text{Forced laborers}_{iy}}{(\text{Forced laborers}_{iy} + \text{True Reservists}_{iy})} \tag{3}$$

This variable is defined for year y, and a given village (locality) i, or a given district d, although it can be observed at the locality level only for a restricted sample of years (1941, 1942, 1944 and 1945).[31]

[30]It's important to note here that volunteers are excluded from this specification, as the district-level lottery system only applies to the second-portion military reservists as well as the major segment of the first-portion regular soldiers (called "appelé" in French colonial military report terms) net of volunteers (called "engagé volontaire" in colonial terms). And we believe that the number of volunteers that a given village supplied for military service could be very endogenous to village-level characteristics which we cannot control, as such they are currently left out of our main specification.

[31]For 1944 and 1945, we do not have individual-level information on forced labor activation from the

4.2 Demographic Burden of Conscription on the Local Population

The military conscription system (together with the forced labor entailed) did pose a significant demographic constraint on the local population in colonial Mali, mobilizing 3% to 10% of the 20-year-old population.[32] In order to assess this in a more rigorous manner, we take the demographic model from Cogneau and Mo (2023), where the authors have simulated population structures for different age groups in all colonies of French West Africa from 1915 until 1960.

Given that the target of military recruitment was 20-year-old young males, by using district-level conscription data collected from the National Archives of Sénégal, we calculate the shares of targeted males under different steps of the conscription system, out of the total projected 20-year-old young males in each single year for the entire colony of Mali.

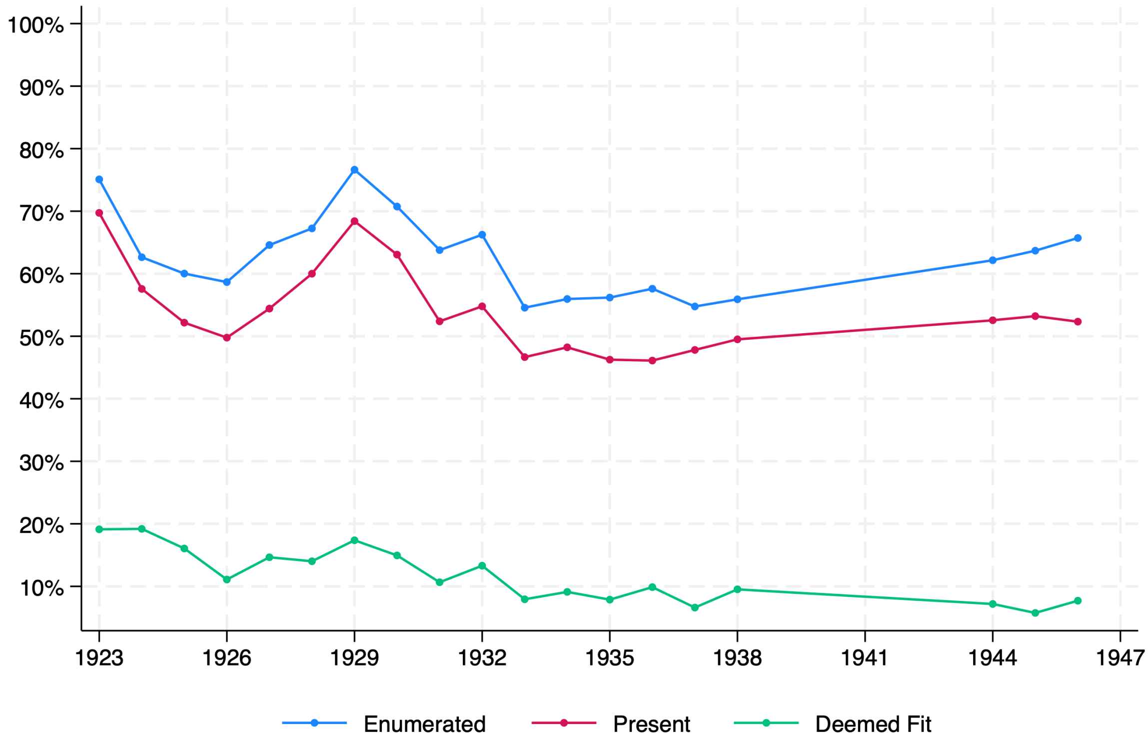
For instance, in Figure 4, it could be observed that the number of enumerated individuals (deemed eligible for military service) stands consistently at around 60 to 70% of the total projected 20-year-olds in colonial Mali across time, with the figure even reaching at around 80% in some years. The present share closely follows the enumerated share, and the gap between the two indicates the absentee share among the enlisted young males. The relatively narrow gap between the two demonstrates that the conscription system was coercive enough in its implementation (to prevent absenteeism). However, the general health conditions of young Africans in colonial years seemed rather deplorable, with the deemed fit young males standing at around only 20% of the total 20-year-olds in the 1920s, and this figure kept decreasing further to around 10% over time.

As previously discussed, among the fit people, around two thirds of them would ultimately become second-portion reservists and the rest would become first-portion regular army recruits. This is observed in Figure 5 where the reservist share always appears to be larger than the regular-soldier share.

CAPM data. However, from the archival records of ANS district-level data, we know that all the reservists recruited in these two years were activated, which means the activation rate was 100%.

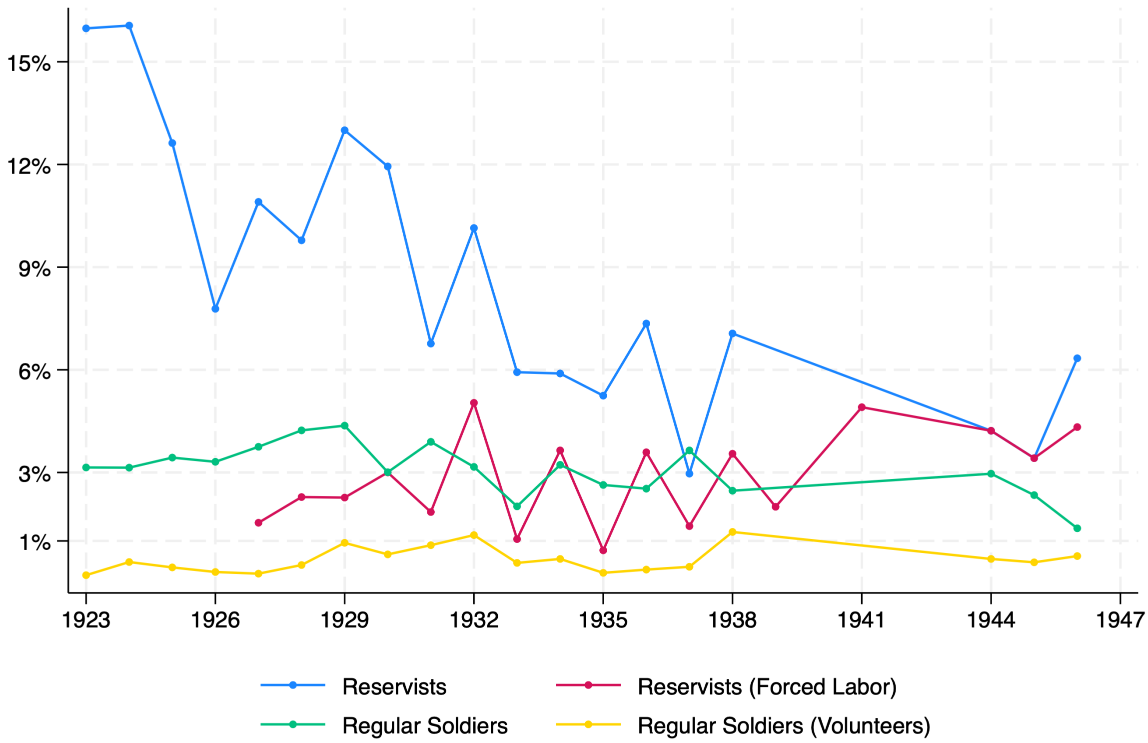
[32]These figures are obtained by summing up the conscript shares in Figure 5.

Figure 4: Shares of Enumerated, Present and Deemed Fit out of Total 20-Year-Old Young Males



Notes: This figure depicts the share of military conscripts by different categories out of the total number of 20-year-olds for the entire colony of Mali annually. The military conscript data comes from ANS, and the demographic projection comes from Cogneau and Mo (2023).

Figure 5: Shares of Soldiers and Reservists out of Total 20-Year-Old Young Males



Notes: This figure depicts the share of military conscripts by different categories out of the total number of 20-year-olds for the entire colony of Mali annually. The military conscript data comes from ANS, and the demographic projection comes from Cogneau and Mo (2023).

4.3 Short-term Reactions towards Military Forced Labor

There is some historical anecdotal evidence indicating that the activation of forced labor prompted young males in colonial Mali to volunteer for the regular army, in order to escape the traumatic experience associated with forced labor.

For instance, one report written by a certain district administrator in colonial Mali reads as the following: "*However, it should be noted that most of the volunteering commitments made in Segou and Bougouni*[33] *were motivated by the fear felt by young people of being classified into* ***"la deuxième portion"*** *and assigned to the Workers' Companies......* ***This is how the district of Segou, out of 280 recruits, has 228 voluntary enlistments."***[34]

We empirically verify this with the district-year-level conscript data digitized from ANS. Specifically, we are interested in that in a given district, whether the forced labor intensity in the previous years resulted in more volunteering into the army in the current year. As such, we regress the total number of volunteer soldiers in year t on the number of forced labor reservists in the previous two years (year $t-1$ and year $t-2$) at the colonial district level. We also control for the contemporaneous total number of forced labor workers to account for auto-correlations in the forced labor variable. Given the fact the number of volunteers was capped in colonial Mali, we also try to control for a few other contemporaneous recruitment variables to tease out the endogenous variations in the number volunteers due to the change of the cap in response to recruitment constraints or military quota specified at the district level in a given year (specifically we control for the contemporaneous military quota of soldiers to be recruited, the number of actual soldiers drafted, as well as the total number of reservists recruited).

The results for this short-term reaction towards the forced labor system are reported in Table 4, where we find a very significant and positive relationship between the total number of reservist forced laborers in the previous two years and the total number of volunteer soldiers in year t. In particular, both the total number of reservist forced laborers in year

[33]Both Segou and Bogouni are colonial district names in Mali.

[34]Annual Military Conscription Report (1931-1932) - Série-4D, Archives Nationale de Sénégal, Dakar.

$t-1$ and year $t-2$ seem to be significantly correlated with the number of volunteers in year t, although the significance on the term in year $t-1$ loses some precision when we include more contemporaneous recruitment controls.

It could be interpreted that when the number of forced laborers increased by 50% two year ago, there was a 5% increase in the total number of volunteers. The results are insignificant between current-year forced laborers and current-year volunteers.[35] Furthermore, if we look at the rate outcomes in Table B2, we could observe that when the forced labor activation rate in year $t-2$ increased by 50 percentage points, the current-year volunteering rate into the army would also significantly increase by around 5 percentage points.[36]

The aforementioned results point to a scenario where the military forced labor system was being evaded by the conscripts themselves. They provide suggestive evidence on the imprint the system could have left on the collective mindset at the time, and pave the ground for understanding the long-run shadow of this oppressive event.[37]

[35]It is important to note here that there is a marginally significant impact of the number of forced laborer in the past year on the number of current-year volunteers as well, although due to inter-temporal correlations the effects are more noisy and smaller.

[36]Volunteering rate is defined as the number of volunteers out of the total regular soldier recruits in a district in a given year. Additionally, we do not find a relationship between forced labor and the number of absentees or absenteeism rates.

[37]Furthermore, we do not think this volunteering reaction towards forced labor would undermine the forced labor system, as previously documented, the number of volunteers were capped in this conscription system and the average volunteering rate during this period of time was only around 15% in colonial Mali, which means that there would still be sufficient variation in the lottery rate to assign which segment of the fit males would join the army as first-portion recruits, and which would be assigned to the reservist camp. We believe there is still sufficient power in this lottery (reservist intensity) measure for us to detect long-term outcomes later on (detailed more in the following section).

Table 4: Correlation between Second Portion Forced Labor and Volunteering Outcomes

VARIABLES	(1) ihs(Volunteers)$_t$	(2) ihs(Volunteers)$_t$	(3) ihs(Volunteers)$_t$
ihs(Forced Laborer)$_t$	-0.0456	-0.0285	-0.0684
	(0.0469)	(0.0478)	(0.0523)
ihs(Forced Laborer)$_{t-1}$	0.0789*	0.0587	0.0860**
	(0.0408)	(0.0377)	(0.0395)
ihs(Forced Laborer)$_{t-2}$	0.1160***	0.1017***	0.0921**
	(0.0375)	(0.0355)	(0.0329)
ihs(Soldiers Drafted)$_t$		-0.3702	-0.4094
		(0.2434)	(0.2657)
ihs(Military Quota)$_t$		0.7250***	0.8130***
		(0.2464)	(0.2635)
ihs(Total Reservists)$_t$			0.2653***
			(0.0694)
Observations	337	336	336
R-squared	0.646	0.661	0.673
District FE	YES	YES	YES
Year FE	YES	YES	YES

Notes: Panel regressions at the colonial district-year level with ANS data. We take the inverse hyperbolic form (ihs) of both the explanatory and dependent variables here. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

4.4 Hypothesized Determinants of Conscription and Forced Labor

Our analysis of the determinants of the intensity of military conscription and the allocation of military conscripts is consistent with a top-down process where a central colonial planner optimized coercion and organization of labor acting upon its knowledge of district-level characteristics but rather limited consideration for within district local characteristics.

At the district level, we study the effect of population size, district area, distance to construction sites and historical experience of colonial conquest [38] on the total number of conscripts and the allocation of these conscripts as soldiers, true reservists, and forced workers to each construction site.

The results displayed in the first panel of table 5 confirm that district-level aggregated conscription outcomes including total number of conscripts, share of soldiers, activation rate and share of forced workers sent to either the irrigation or the railways sites depended on labour needs and ease of recruitment.

The total number of conscripts directly reflect the intensity of the military conscription procedure. As documented by historical evidence, this figure would depend on the extent to which customary leaders reported faithfully young men eligible for conscription, whether these young men would present themselves to the conscription center and whether these would be deemed as healthy enough for military service. These steps of the recruitment process are described in the panel (a) of Figure A2 in Appendix.

The colonial authorities recruited more in more populated districts as of the last census of 1925, with an average of 2.7 conscripts per thousands inhabitants, and, controlling for population, recruited less in larger districts and districts conquered later by the French army in the 19th century, both variables indicating a lower ease of access and control of the territory. The expected effect of the distance to railways on intensity of recruitment is a priori ambiguous. On one hand, railways represented about a quarter of the demand

[38]We look at the effect of the year of start of conquest by the French Army, between 1855 and 1896. This data is taken from Huillery (2009)

for forced labor and facilitated the transportation of soldiers and forced laborers, factors that should logically increase the ability and incentives of colonial authorities to recruit more from districts closer to railways. However, as documented by the work of Pruett (2024), railways, by expanding economic opportunities and facilitating mobility, also empowered the population, increasing their ability and incentives to evade conscription. As population at the district level dates is measured only once shortly before our period of interest 1925, it is also possible that railways-induced population growth is not controlled for. As such, the net positive effect of proximity to railways on total number of conscripted we estimate may result from the positive effect of this infrastructure on labour demand and ease of transportation, but it may also be reflective of growing or healthier population in these districts. The negative effect of distance to railways slightly decays with distance as indicated by the positive but small coefficient estimated for the distance squared [39]. Distance to the main construction site, the *Markala* irrigation dam at the *Office du Niger* on the other hand, should only raise the incentives of the colonial authorities to increase total recruitment, as the majority of forced workers were dispatched to the irrigation project (see Table B1. We estimate only a small negative effect of the squared distance to the *Markala* dam on the total number of conscripts and no significant effect of linear distance, suggesting that demand for forced labour concerns played a marginal role at this stage of the conscription process.

The next stage of the recruitment process involved allocating conscripts to the army or reservists, where districts administrators and the FWA General Governor set the lottery rate for each district. Thise steps of the recruitment process ise described in the panel (b) of Figure A2 in Appendix. This step reflects both an allocation decision, determining how recruited individuals are utilized, and signals the intensity of military conscription. The main objective of the colonial authority remained to reach soldiers recruitment target, forced laborers representing only a by-product of this process. The effect of population size and area are similar than those observed for the total number of conscripts, with more populated districts allocating a larger share of their conscripts to military service

[39] The negative effect dominates for distances inferior to 595 000 km

and larger districts presenting a lower share of drafted soldiers. There is no significant effect of the historical year of start of conquest of a district on its share of soldiers. If the district administrators accounted for labour needs when setting the lottery rates, the share of soldiers should be negatively correlated with distance to construction sites. This is not observed in our results, suggesting that demand for forced labour was accounted for only in the last stages of the conscription process.

The rate of activation of reservists as forced workers plays a crucial role in determining the labor supply available for forced labor and represents a significant source of variation in the coerciveness of the conscription system. This activation rate is positively correlated with population size, negatively correlated with district area, but shows no correlation with the historical year of conquest. Proximity to the main irrigation construction site, the *Markala* dam, emerges as a strong determinant of the activation rate. Furthermore, no significant decay of this relationship can be observed with the squared transformation of this distance. In contrast, the distance to railways shows a positive correlation with activation, suggesting that concerns for evasion of forced labor were greater in areas with access to railways.

Finally, the allocation of forced workers to either irrigation or railway construction sites sheds light on the colonial authority's utilization of labor extracted from the local population [40]. This allocation demonstrates no correlation with population size, area, or historical conquest year. Instead, its correlation with proximity to construction sites indicates a strategic and efficient allocation of forced labor. Districts located nearer to irrigation sites tend to dispatch more labor to these projects and less to railway construction, whereas districts in proximity to the railway line exhibit the opposite pattern. Note that the effect of distance to railways is weaker, implying that the proximity to the *Kayes-Bamako* railways imperfectly captures the exact location of the construction sites along the line.

In the second part of Table 5, we reproduce the same analysis at the locality level using the CAPM data. We focus on the solider-share and activation variables, by definition scaled

[40] Columns (4) and (5) are perfectly symmetric as share irrigation = 1 - share railways

to the total number of conscripts, as population data at the locality level are not available for the colonial period.

In the absence of district fixed effects (columns 1 and 4), the findings remain largely consistent with the district-level results. Localities in closer proximity to the *Markala* dam tend to exhibit lower shares of soldiers (higher share of reservists), as well as higher activation rates, while localities near railways have on average a lower share of soldiers (higher share of reservists) but lower activation rate of reservists as forced workers. The correlation between distance to the *Markala* irrigation dam and the share of soldiers is stronger at the locality level as both the linear and square distance have a positive significant effect, which could be attributed to the increased granularity of the distance variable at the locality level.

Once controlling for district fixed effects, however, within-district variation in distance to construction sites does not appear to strongly influence locality-level conscription outcomes. As expected, the share of soldiers, determined by a lottery draw at the locality level, shows no correlation with any of the distance variables. The determinants of the activation rate at the locality level can only be estimated for a sub-sample of localities with military recruitment in 1941 and 1942. In these years, the activation rate was higher than the period average and, in some districts, close to one hundred percent, resulting in smaller variation. Even within this smaller sub-sample, a significant correlation is observed between within-district variation in activation rate and proximity to the *Markala* irrigation dam. This suggests that the determination of which reservists would be dispatched as forced laborers was a more decentralized process than the lottery.

Table 5: District and locality-level determinants of recruitment of reservists and soldiers

	District-year level ANS data				
	Concripts	Share soldiers	Activation	Irrigation	Railways
	(1)	(2)	(3)	(4)	(5)
Pop. 1925	2.757***	0.0003*	0.001**	0.001	−0.001
	(0.254)	(0.0002)	(0.0004)	(0.0004)	(0.0004)
Area	−248.607*	−0.136**	−0.151**	−0.056	0.056
	(143.814)	(0.069)	(0.065)	(0.210)	(0.210)
Dist. Markala	0.227	−0.0002	−0.001**	−0.002***	0.002***
	(0.500)	(0.0003)	(0.001)	(0.001)	(0.001)
Dist. Markala sq.	−0.002*	0.00000*	0.00000***	0.00000	−0.00000
	(0.001)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Dist. railways	−1.184***	0.001***	0.001*	0.001*	−0.001*
	(0.221)	(0.0002)	(0.0005)	(0.0005)	(0.0005)
Dist railways sq.	0.002***	−0.00000***	−0.00000**	0.00000	−0.00000
	(0.0004)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Conquest start year	−10.234**	0.001	0.005	−0.002	0.002
	(4.700)	(0.002)	(0.003)	(0.007)	(0.007)
Mean DV	382.68	0.32	0.53	0.70	0.30
Year FE	Yes	Yes	Yes	Yes	Yes
Obs	348	347	339	172	172
Adj. R^2	0.480	0.345	0.497	0.565	0.565

	Locality-year level CAPM data			
	Share soldiers		Activation	
	(1)	(2)	(3)	(4)
Dist. Markala	0.0001**	0.00004	−0.001***	−0.0004
	(0.00004)	(0.0001)	(0.0001)	(0.0002)
Dist. Markala sq.	0.00000*	0.00000	0.00000***	0.00000*
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Dist. railways	0.0001*	0.00003	0.002***	0.0003
	(0.00003)	(0.00005)	(0.0001)	(0.0002)
Dist. railways sq.	−0.00000**	−0.00000	−0.00000***	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Mean DV	0.28	0.28	0.73	0.73
Year FE	Yes	Yes	Yes	Yes
Cercle FE	Yes	Yes	Yes	Yes
Obs	56,590	56,590	4,317	4,317
Adj. R^2	0.094	0.106	0.094	0.198

Notes: We include year fixed effects in all regressions and colonial district fixed effects in the locality-year level regressions. District level data is available from aggregate sources for all years but 1939 to 1944 and 1948 to 1950. Locality level data is available for all years excepting the distribution of reservists between true reservists and forced workers (i.e. the activation rate) available for 1941 and 1942 only. Standard errors clustered at the circle level for the circle level analysis and at the locality level for the locality-level analysis. Population data at the district level comes from an AOF census collected in 1925 in the whole French West Africa digitized by (Huillery, 2009). Distance to the railways is proxied by distance to a straight line between Bamako and Kayes. Significance levels: *p<0.1; **p<0.05; ***p<0.01

It does not come as a surprise that the share of reservist is independent of locality-level proximity to construction sites, given historical evidence that if there were any strategic concerns on the recruitment and allocation of drafted young males, they should have taken place already at the colonial district level, but shouldn't have boiled down to the locality level. For the forced labor activation measure, while less is known about how the fraction of reservists activated as forced labor workers was distributed within districts, the centralized set-up of the drafting process within districts could be compatible with a distribution orthogonal to local conditions.

Tables B6 in Appendix supports the exogeneity of the locality-level soldier-share variable. Table B6 documents the correlation between ethnic composition of localities, as measured with the 2009 population census data, and military conscription outcomes. While the activation rate and number of volunteers seems to be strongly correlated with the ethnic composition inside a given locality in colonial Mali, even inside a locality within a given district, the share of reservists (soldiers) is only very weakly correlated with the share of the main ethnic groups after controlling for colonial district-year fixed effects, and such correlation is significant at the 1% level only for one ethnic group out of ten.

While further analysis should be conducted to establish the exogeneity of within-district variations in the assignment of conscripted individuals at the locality level, these results suggest that once the enumerated 20-year-olds had passed medical exams and been conscripted into the army, their assignment as either soldiers or reservists was strongly correlated with demands and constraints in their districts of recruitment, but irrespective of their local conditions at the village level.

These historical determinants that shaped military recruitment at the district level motivate our empirical strategy to exploit within-district, rather than between-district variation in the share of soldiers, and to consider relative measures of forced labor (i.e. the share of reservists among total conscripts and the fraction of reservists activated as forced labor) rather than absolute numbers of individuals recruited as forced workers.

5 Long-term Persistence of Forced Labor on Development Outcomes

5.1 Empirical Strategy

A quasi-random proxy of exposure to forced labor: locality-level share of soldiers To assess the relationship between forced labor and development outcomes over the long run, as discussed in Section 4.1, we construct a cumulative share of soldiers variable at the village level, as the ratio of total drafted soldiers recruited in the village to the total number of conscripts of the village over the full 1927-1950 period, during the period of which forced labor was in place as specified in equation 4:

$$\text{Share\_Soldiers}_i = \frac{\sum_{y=1927}^{1950} \text{Drafted Soldiers}_{iy}}{\sum_{y=1927}^{1950} (\text{Reservists}_{iy} + \text{Drafted Soldiers}_{iy})} \tag{4}$$

for year y, in a given village (locality) i.[41]

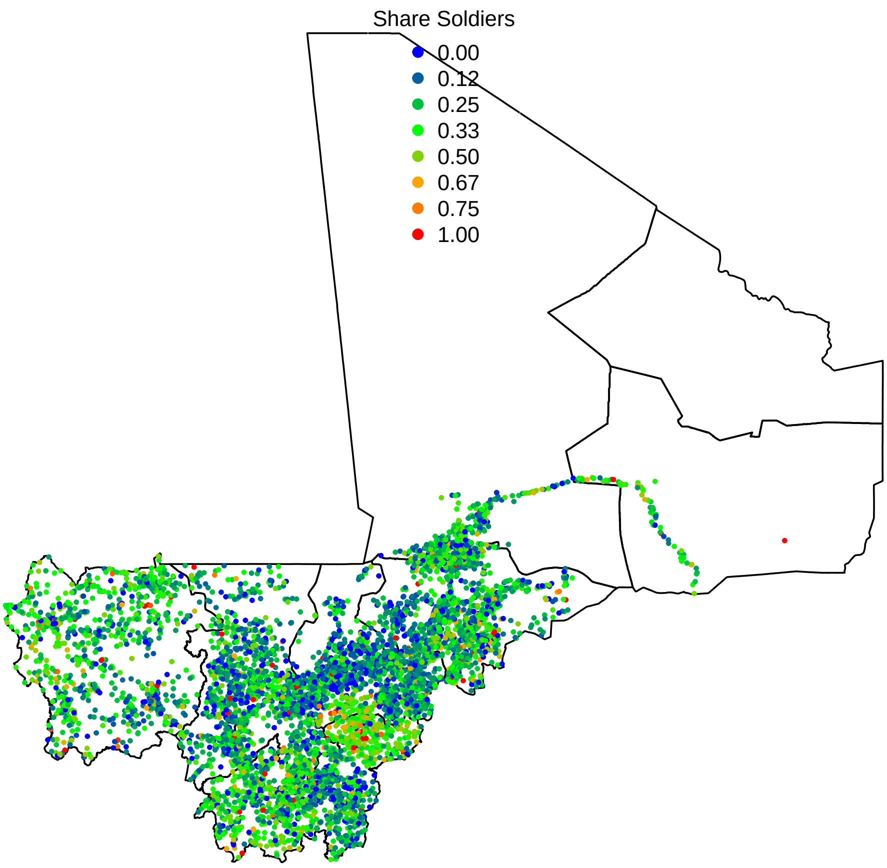
Figure 6 shows the spatial distribution of the soldier share variable across localities in Mali. While on the one hand we could observe that there is considerable variate of the share of soldiers across localities, on the other hand, some localities with low share of soldiers are also clustered in certain colonial districts.

As discussed in Section 2, at the colonial district level, the share-of-soldier measure described above is endogenous given that for a given year, this share[42] was adjusted locally by the district administrator to fulfil the military recruitment quota assigned to that district by the Lieutenant Governor of colonial Mali (see Figure A2a), depending on conditions of local labor scarcity and the ease of recruitment. As these characteristics may

[41]It's important to note here that volunteers are excluded from this specification, as the district-level lottery system only applies to the second-portion military reservists as well as the major segment of the first-portion regular soldiers (called "appelé" in French colonial military report terms) net of volunteers (called "engagé volontaire" in colonial terms). And we believe that the number of volunteers that a given village supplied for military service could be very endogenous to village-level characteristics which we cannot control, as such they are currently left out of our main specification.

[42]It's previously labelled as "lottery rate".

Figure 6: Geographical Distribution of Locality-level Share of soldiers for matched localities in the 1976 census.



correlate with long-term development outcomes, we exploit the within-district variation in the share of soldiers (out of total conscripts). The within-district variation was at the time, in theory, independent of village-level characteristics, as all drafted young males were assigned a number drawn from a lottery that would determine whether they would serve as soldiers or reservists.[43] Given that each village sent on average about 20 individuals over the whole 1927-1950 period, in the case of which we are dealing with an effective small sample at the village level, cumulatively speaking the village share of soldiers does not converge to the district-level share of soldiers. Intuitively, the within-district locality-level variation in soldier share that we rely on for identification is similar to a sampling error generated by small sample size. Monte-Carlo simulations described in section A.8 in Appendix confirm that the observed variation in locality-level share of reservists (soldiers) falls within the range of randomly generated distributions.

Another test for the randomness of the lottery rate is performed in Table B4: after controlling for district-by-year fixed effects, there seem to be no serial correlation of the share of reservists (soldiers) over periods. This means that a locality with a higher share of reservists relative to the average district share in a given period will not necessarily display a higher share in the following period. If the lottery rate would have been determined by some unobserved structural characteristics of localities, such a serial correlation would have been detected, so that time-in-varying omitted variables are not likely to drive the results. Furthermore, Table B5 shows that within district, there is no correlation between a locality share of soldiers and the average share of soldiers of other localities within a radius of 5 km, whereas such spatial correlation can be detected for the share of volunteers. This further aligns with the historical analysis in Table 5 that the lottery rate was orthogonal to locality-level distance to construction sites that were otherwise relevant in determining district-level lottery rate and locality-level activation rate.

Although there is some anecdotal evidence indicating that within a given village, lottery outcomes might still be manipulated, these suggest rather a within-locality manipulation than cross-locality. For instance chiefs' sons would predominantly serve as soldiers,

[43]The whole procedure was described in detail in Section 2.

rather than reservists, and local elites would send lower ranked individuals from their villages in case of a bad draw.[44] While this manipulation could affect the distributional impacts of exposure to conscription, cross-locality variation in exposure would remain exogenous. Further empirical tests should help us determine the extent of manipulation across villages. Nevertheless, we believe that the simulation exercise and our tests largely supports the quasi-random nature of our primary explanatory variable.

The econometric model of our identification strategy is specified as the following:

$$Y_{iy} = \beta_0 + \beta_1 \text{Share\_Soldiers}_i + \beta_c + u_i \tag{5}$$

where the unit of analysis is a locality i in a census year y, Y_{iy} is the outcome of interest, β_c controls for district fixed effects.

Estimation of equation 5 indicates the overall effect of being randomly assigned to either the first-portion regular soldier army, compared to being assigned to the military reservist camp (hence higher likelihood of being activated as forced labor workers) on contemporary development outcomes. Albeit being able to identify long-term impacts of being a soldier relative to a reservist in a causal framework, this estimation cannot tell us the relative impacts of being a forced labor worker, relative to being soldiers or true military reservists. With this shortcoming in mind, in order to better disentangle the effects of being a true military reservist relative to being a forced labor worker, we exploit another variable at the colonial district level, which is the **activation of reservists as forced workers**. We define the following district level activation rate over the full period as the following:

$$\text{Activation}_d = \frac{\sum_{y=1927}^{1950} \text{Forced laborers}_d}{\sum_{y=1927}^{1950} (\text{Forced laborers}_d + \text{True Reservists}_d)} \tag{6}$$

[44] We find in our archival data that chiefs' sons were selectively recorded on separate sheets with their personal information (such as names and village of origins, etc), as well as the nature of their military drafts. The overwhelming majority of them were first-portion regular soldiers. These data will be exploited more extensively in the future.

This measure, as discussed in Section 4, was endogenously determined by the colonial administration and was correlated with district level characteristics such as population and distance to public work construction sites. It is likely to be correlated with district-level unobservable characteristics as well. For these reasons, we do not consider directly the effect of this activation rate, but rather compare the effect of the reservist share measure in localities with low and high fractions of forced workers among reservists across colonial districts. The spatial distribution of the forced labor activation rates can be seen in Figure 7.

The corresponding estimation equation is as the following:

$$Y_{iy} = \beta_0 + \beta_1 \text{Share\_Soldiers}_i + \beta_2 \text{Activation}_d + \beta_3 \text{Share\_Reservists}_i * \text{Activation}_d + \beta_c + u_i \tag{7}$$

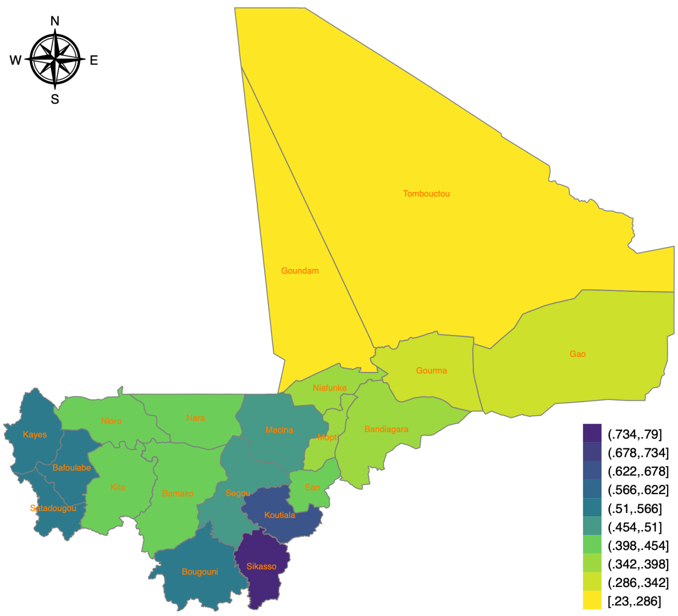
where the unit of analysis is a locality i in a census year y, Y_{iy} is the outcome of interest, β_c controls for district fixed effects, and $Share\_Reservists_i$ is equal to $1 - Share\_Soldiers_i$.

In this estimation, the coefficient β_1 on the share-of-soldiers term alone should capture the effect of the share of soldiers when the activation rate is null, so that it captures the effect of young men being drafted as regular soldiers compared to being true reservists (not activated as forced labor). Similarly, the intercept coefficient for activation β_2 should capture the effect of activation when the share of reservists is null. As this share of activated forced workers among reservists should not affect development outcomes in localities with a null share of reservists, this coefficient captures the omitted variable bias associated with determinants of activation rate at the district level.[45] Finally, the main coefficient of interest is the interaction between activation and the share of reservists, which is our proxy for forced labor exposure. Namely, β_3 captures the additional effect of the reservist-share variable as the fraction of forced workers within reservists increases. In this scenario, our empirical identification could be rationalized as a modified form

[45] Although this term is automatically controlled for when we control for district fixed effects.

of continuous Differences-in-Differences strategy. In particular, the first difference is the more exogenous variable, namely the differing share of reservists (out of total recruits) across different localities. The second difference, is the varying forced-labor activation rate across colonial districts. For the interaction term between the two, β_3, to identify the causal effect of forced labor, we need the following "parallel trend" assumption. In our context, this assumption is that in the absence of forced labor activation, the effect of the share of reservists should be the same between high-activation districts and low-activation districts, so that the heterogeneity of the effect is only driven by forced labour. While further tests should document the validity of this assumption, historical evidence suggest that there should not be any difference in the treatment of soldiers and reservists across districts other than the share of activation of reservists as forced labour as soldiers were assigned to different battalions irrespective of their district of origin (see Salem and Seck (2022) for empirical tests in Morocco).

Figure 7: District-Level Military Forced Labor Activation Rate (Out of Total Reservists)



Notes: This figure depicts the cumulative average activation rate of forced labor workers among total reservists at the colonial district level for the time period 1927-1950.

5.2 Results

As we begin to assess the relationship between military forced labor and development, it is worth restating again how we hypothesize that the intensity of forced labor will impact long-term development outcomes in Mali.

First of all, given the relatively small demographic shock the conscription and forced labor system exerted on the local population, we do not expect the long-term persistent effects of military forced labor to go through demographic channels. That being said, we assess if there's any bearing of exposure to military forced labor contemporary demographic outcomes given that the shock was nonetheless still an important in-kind tax on the most able-bodied prime-age labor force at the locality level.

Instead, we argue that the effects of exposure to forced labor are more likely to operate through the inter-generational transmission of values and norms. Specifically, we hypothesize that exposure to colonial military forced labor could result in the rejection of colonial institutions or those that evoke colonial setups among indigenous populations. One such institution is the modern schooling system, which was heavily promoted by the French colonial authorities throughout their African colonies. As stated in the 1987 census report from the National Statistics Institute of Mali, *"Instruction was one of the essential instruments of colonial domination, whose ultimate goal was the moral and intellectual assimilation of peoples."*. Although there has been substantial efforts to reform the schooling system after independence and a progressive extension of bilingual education since 1984, French remains the primary teaching language and formal schooling may still be perceived as another French legacy. As such, the long-term impact of forced labor could be manifested in worse human capital accumulation outcomes, and hence potentially worse overall development outcomes such as wealth, infrastructure presence, etc at the locality level. When it comes to the expected effect on inequality, it is more ambiguous as the potentially negative forced labor shock affected primarily the healthiest individuals, which could be inequality-reducing. However, there is some anecdotal historical evidence suggesting that individuals ultimately affected by the conscription system might not have been the most privileged ones in colonial times (Mann, 2006), hence pointing towards a potentially inequality-increasing effect of forced labor as well.

5.2.1 Main results on development outcomes

The results are reported in Table 6. Results for the long-term effect of the share of soldiers among drafted conscripts (estimating equation 4) are displayed in the odd columns and the heterogeneous effects of this lottery rate estimated with the interaction between the share of reservists and forced labor activation rate (estimating equation 7) are shown in the even columns.

First of all, in terms of the effects of the share of soldiers alone, consistently across different waves of censuses, we do not detect a significant impact of this village-level lottery rate into the army on contemporary development outcomes, either in terms of fertility levels or the share of villagers having attained at least primary schooling, with the exception of a weakly significant positive coefficient for primary schooling in 2009. The effects are also null on village-level mean wealth indices and wealth inequality measures in the 1987, 1998, and 2009 censuses.

However, such an average null effect conceals significant heterogeneous effects of the reservist-share variable depending on the activation rate of reservists as forced workers across districts.

In terms of the human capital outcomes, we find that for null activation rates, meaning in the absence of forced labor, the long-term effect of a larger share of men being drafted as soldiers (reservists) relative to reservists (soldiers) is negative (positive). This effect is highly significant across all four waves of Malian census. In districts with null activation rates (activation = 0), a one percentage point increase in the locality share of soldiers is equivalent to a one percentage point decrease in the share of true reservists, relative to soldiers, and resulted in a 0.08 percentage points reduction in the share of individuals of this locality who attended at least the primary school level in 1976. It indicates that there might be a negative shock associated with being enlisted in the French army (such as

Table 6: Effect of military service and forced labor on development outcomes

	Census 1976							
	Fertility		Primary school		Eligible Men Pop			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.00	0.11**	0.00	-0.08***	0.59	-42.91		
	(0.01)	(0.06)	(0.01)	(0.03)	(9.74)	(52.73)		
Share_Reservists x District Activation Rate		0.25**		-0.18***		-96.02		
		(0.12)		(0.07)		(118.74)		
Mean DV	0.44	0.44	0.08	0.08	99.53	99.53		
St. Dev. DV	0.14	0.14	0.08	0.08	146.95	146.95		
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes		
Obs	4271	4271	4276	4276	4291	4291		
Adj. R^2	0.12	0.12	0.08	0.08	0.08	0.08		

	Census 1987							
	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.09	0.00	-0.06***	-0.03	-0.10	-0.00	-0.07
	(0.02)	(0.07)	(0.01)	(0.02)	(0.03)	(0.13)	(0.03)	(0.14)
Share_Reservists x District Activation Rate		0.16		-0.13***		-0.14		-0.15
		(0.14)		(0.05)		(0.27)		(0.28)
Mean DV	0.88	0.88	0.06	0.06	-0.15	-0.15	0.34	0.34
St. Dev. DV	0.22	0.22	0.08	0.08	0.43	0.43	0.39	0.39
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5607	5607
Adj. R^2	0.17	0.17	0.08	0.08	0.26	0.26	0.09	0.09

	Census 1998							
	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.07	-0.00	-0.09***	0.01	0.02	0.02	-0.04
	(0.01)	(0.06)	(0.01)	(0.03)	(0.02)	(0.08)	(0.02)	(0.10)
Share_Reservists x District Activation Rate		0.13		-0.18***		0.02		-0.14
		(0.12)		(0.06)		(0.18)		(0.22)
Mean DV	0.82	0.82	0.09	0.09	-0.31	-0.31	0.17	0.17
St. Dev. DV	0.18	0.18	0.11	0.11	0.26	0.26	0.31	0.31
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5607	5607
Adj. R^2	0.10	0.10	0.14	0.14	0.13	0.13	0.12	0.12

	Census 2009							
	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.04	-0.02*	-0.14***	-0.01	-0.01	-0.00	-0.03
	(0.02)	(0.08)	(0.01)	(0.05)	(0.02)	(0.09)	(0.02)	(0.08)
Share_Reservists x District Activation Rate		0.06		-0.25**		0.02		-0.07
		(0.15)		(0.10)		(0.20)		(0.17)
Mean DV	0.94	0.94	0.23	0.23	-0.41	-0.41	0.29	0.29
St. Dev. DV	0.27	0.27	0.16	0.16	0.32	0.32	0.29	0.29
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5605	5605	5607	5607	5607	5607	5607	5607
Adj. R^2	0.02	0.02	0.23	0.24	0.24	0.24	0.21	0.21

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. For census 1976: Schools, Dispensaries, Markets indicate the presence of the infrastructure in the locality. For census 1987 and 2009: Fertility is measured as the ratio of children between 0 and 5 over the number of women 15-50 in the village. Primary school is measured as the share of the population in the locality aged 6 years or more, having attained at least primary school; Wealth Index is the average of the wealth index computed using principal component analysis of the housing characteristics of the households within the locality; Interquartile range is the difference between the wealth index of the 10 percentile and the 90th percentile in the village. District fixed effects are colonial fixed effects. Discrepancies in the number of observations for Gini is due to insufficient number of households in the locality to calculate Gini. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01

high mortality rates associated with World War Two service, etc) compared to not having been exposed to any conscription labor shock. When the force labor activation rate turns positive, we find a reversal of the effect of the lottery rate due to the additional negative impact of exposure to military forced labor. In other words, in districts with high activation rates, localities with higher share of soldiers were ***relatively*** better off than localities with high share of reservists, since a high share of young men enrolled in the army for all its potential negative impacts, also meant a high share of young men avoiding forced labor. In districts with full activation rates (activation = 1), a one percentage point increase in the locality share of reservists is equivalent to a one percentage point increase in the share of forced workers, relative to soldiers, and resulted in a 0.10 [46] percentage points reduction in the share of individuals of this locality who attended at least the level primary school in 1976. This implies that, in the average village, a one percentage point increase in the share of forced workers, relative to true reservists, resulted in a 0.18 percentage points decrease in the share of ever educated individuals in 1976 [47]. Hence the negative long-term effect of forced labor is twice as large as the negative long-term effect of military service.

However, we can detect no significant effect on wealth index or inequality indicator, which is rather surprising given the magnitude of the effects on education, and would suggest low returns to schooling in Mali[48].

Ultimately, with regards to the demographic effects, there appears to be a positive effect of both military service and forced labour on fertility outcomes in 1976. Further investigation on mechanisms is needed to better understand the effects detected on fertility in the early years after independence, although the effect turn insignificant in the subsequent waves of censuses. Furthermore, in the 1976 census, we also attempt to assess if

[46]We subtract the positive (negative) effect of share reservists (soldiers) without null activation: 0.18 - 0.08 = 0.10.

[47]Another way to read the interaction term is: comparing villages with the same share of reservists in districts where the activation rate of forced workers is either 0 or 1, villages in districts with full activation will have on average18 percentage points lower share of individuals who ever attended at least primary school in 1976 than those in districts with no activation of reservists.

[48]It is also possible that our measure of wealth, building an index with household assets, is too noisy to capture these returns.

exposure to forced labor impacted the total number of male population who were eligible for military service (hence forced labor) during the colonial years. Although the immediate demographic effect is not expected to be large given the total recruitment rates, it could be further amplified by migration responses to conscription pressure. We identify the individuals who were born between the years of 1907 and 1930, hence they were of 20 years old in the forced labor period 1927-1950. However, the demographic effects on them appear to be null, which lends evidence to the claim that the exposure to military conscription did not result in a persistent demographic shock at the locality level.

We reproduce the main results excluding the North districts which differ from the average districts in terms of matching rate and activation rates in Appendix in Table C6 : the coefficients for primary school are of slightly larger magnitude and more precisely estimated.

5.2.2 Infrastructure and supply effects

One question that arises from the long term results on school attainment is whether the long-run decrease in human capital outcomes is more driven by a supply effect or demand effect. To better understand the specific channels, we exploit both the infrastructure variables in the 1976 census (number of schools, dispensaries, etc) and the most recent exhaustive census of all (primary and secondary) schools constructed in Mali up until April, 2021.[49] In Table 9, using the same empirical specification as in Table 6, we observe that in the first panel using the 1976 Malian census, the results already hint at a potential supply effect: both exposure to regular military service and military forced labor decrease the total number of schools at the locality level, with the effect of the latter being more negative. Nonetheless, this effect is only marginally significant and the measurement is also quite noisy in the earliest wave of the post-colonial census.

When we move on to the second panel using the school census in 2021, there is no bearing of our variable of interest, the locality-level share of reservists interacted with the district-level forced labor activation rate, on the total number of different types of schools, as well

[49]This school census is part of the outcome resulting from the 2021 population census.

Table 7: Effect of military service and forced labor locality infrastructure presence

	Infrastructure Census 1976							
	Schools		Health		Pharma_Health		Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.01	-0.16*	0.01	-0.06	0.02	-0.07	0.00	0.03
	(0.02)	(0.08)	(0.02)	(0.09)	(0.03)	(0.11)	(0.00)	(0.03)
Share_Reservists x District Activation Rate		-0.39*		-0.20		-0.25		0.07
		(0.21)		(0.24)		(0.30)		(0.08)
Mean DV	0.11	0.11	0.09	0.09	0.11	0.11	0.01	0.01
St. Dev. DV	0.32	0.32	0.38	0.38	0.50	0.50	0.19	0.19
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5607	5607
Adj. R^2	0.03	0.03	0.01	0.01	0.01	0.01	0.01	0.01

	Infrastructure Census 1987							
	Schools		Health		Pharma_Health		Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.02	-0.24*	-0.00	-0.11	-0.00	-0.12	0.01	-0.00
	(0.03)	(0.13)	(0.02)	(0.08)	(0.03)	(0.12)	(0.01)	(0.06)
Share_Reservists x District Activation Rate		-0.70**		-0.29		-0.32		-0.03
		(0.34)		(0.20)		(0.31)		(0.18)
Mean DV	0.16	0.16	0.09	0.09	0.14	0.14	0.02	0.02
St. Dev. DV	0.50	0.50	0.33	0.33	0.50	0.50	0.26	0.26
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5607	5607
Adj. R^2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

	Infrastructure Census 1998							
	Schools		Health		Pharma_Health		Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.01	-0.49*	0.01	-0.47***	0.03	-0.54**	-0.13	-0.47
	(0.07)	(0.27)	(0.04)	(0.17)	(0.06)	(0.25)	(0.16)	(0.64)
Share_Reservists x District Activation Rate		-1.28*		-1.29***		-1.53**		-0.91
		(0.71)		(0.45)		(0.66)		(1.60)
Mean DV	0.42	0.42	0.27	0.27	0.39	0.39	1.80	1.80
St. Dev. DV	0.95	0.95	0.73	0.73	1.06	1.06	2.87	2.87
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5581	5581	5581	5581	5581	5581	5581	5581
Adj. R^2	0.03	0.03	0.03	0.03	0.03	0.03	0.06	0.06

	Infrastructure Census 2009							
	Schools		Health		Pharma_Health		Water	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.16	-0.79	-0.01	-0.40***	0.03	-0.50**	-0.10	-0.49
	(0.17)	(0.59)	(0.04)	(0.15)	(0.06)	(0.24)	(0.15)	(0.62)
Share_Reservists x District Activation Rate		-1.70		-1.05***		-1.43**		-1.04
		(1.31)		(0.40)		(0.59)		(1.59)
Mean DV	1.01	1.01	0.23	0.23	0.30	0.30	2.05	2.05
St. Dev. DV	1.72	1.72	0.49	0.49	0.76	0.76	2.48	2.48
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5374	5374	5374	5374	5374	5374	5374	5374
Adj. R^2	0.04	0.04	0.06	0.06	0.04	0.04	0.09	0.09

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. District fixed effects are colonial fixed effects. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table 8: Effect of military service and forced labor locality health infrastructure presence

	CSCOM		Pharmacy		Clinic		Dispensary		Maternity		Hospital	
Health Infrastructure Census 1998	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share_Soldiers	0.00	-0.03	0.03	-0.54**	0.00	-0.04	0.01	-0.13*	-0.00	-0.25***	0.00	-0.02
	(0.01)	(0.03)	(0.06)	(0.25)	(0.01)	(0.03)	(0.02)	(0.08)	(0.02)	(0.09)	(0.00)	(0.02)
Share_Reservists x District Activation Rate		-0.08		-1.53**		-0.12		-0.38*		-0.67***		-0.05
		(0.09)		(0.66)		(0.08)		(0.20)		(0.25)		(0.04)
Mean DV	0.03	0.03	0.39	0.39	0.00	0.00	0.11	0.11	0.13	0.13	0.00	0.00
St. Dev. DV	0.16	0.16	1.06	1.06	0.06	0.06	0.32	0.32	0.39	0.39	0.05	0.05
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5581	5581	5581	5581	5581	5581	5581	5581	5581	5581	5581	5581
Adj. R^2	0.01	0.01	0.03	0.03	0.00	0.00	0.02	0.02	0.03	0.03	0.01	0.01

	CSCOM		Pharmacy		Clinic		Dispensary		Maternity		Hospital	
Health Infrastructure Census 2009	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share_Soldiers	-0.02	-0.16	0.03	-0.50**	-0.00	-0.05	-0.00	-0.03	0.01	-0.13	0.00	0.01
	(0.02)	(0.10)	(0.06)	(0.24)	(0.01)	(0.04)	(0.01)	(0.05)	(0.02)	(0.08)	(0.00)	(0.01)
Share_Reservists x District Activation Rate		-0.39		-1.43**		-0.13		-0.08		-0.40*		0.03
		(0.27)		(0.59)		(0.10)		(0.13)		(0.23)		(0.03)
Mean DV	0.14	0.14	0.30	0.30	0.01	0.01	0.03	0.03	0.05	0.05	0.00	0.00
St. Dev. DV	0.36	0.36	0.76	0.76	0.12	0.12	0.17	0.17	0.23	0.23	0.05	0.05
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5374	5374	5374	5374	5374	5374	5374	5374	5374	5374	5374	5374
Adj. R^2	0.02	0.02	0.04	0.04	0.00	0.00	0.02	0.02	0.07	0.07	0.00	0.00

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. District fixed effects are colonial fixed effects. Standard errors clustered at the locality levels for all regressions. CSCOM refers to Centre de Santé Communautaire, which groups together different kinds of health facilities from the mid-1990s onwards. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table 9: Effect of military service and forced labor on local school supply

School Census 2021	Any School		No. of Schools						Student/Teacher Ratio	
			Community		Quranic		Public			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share_Soldiers	-0.05	-0.01	-0.00	0.05	-0.09	0.16	-0.11	0.62	-0.55	-15.61**
	(0.04)	(0.16)	(0.07)	(0.21)	(0.13)	(0.37)	(0.21)	(1.09)	(1.57)	(7.09)
Share_Reservists x District Activation Rate		0.07		0.12		0.51		1.54		-30.23**
		(0.30)		(0.41)		(0.54)		(1.99)		(13.66)
Mean DV	0.66	0.66	0.33	0.33	0.18	0.18	0.77	0.77	40.82	40.82
St. Dev. DV	0.47	0.47	0.75	0.75	1.61	1.61	2.36	2.36	17.64	17.64
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5607	5607	3706	3706
Adj. R^2	0.13	0.13	0.07	0.07	0.01	0.01	0.02	0.02	0.10	0.10

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. District fixed effects are colonial fixed effects. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01

as the incidence of school existence in Malian villages today.[50] However, there is a statistically significant effect on the student-to-teacher ratio,[51] where we detect a significant negative effect of both exposure to military service and forced labor on the number of students per teacher in a given locality. If we compare two localities with the same reservist share, for the one located in a former colonial district with full activation of reservists as forced labor, the number of students per teacher will decrease by 30 compared to a village in a district with no activation of reservists, which is nearly 80% of the national locality-level mean.

Based on these results above, we currently tend to conclude that the worsening human capital outcomes in localities more exposed to colonial military forced labor are more driven by demand effects rather than supply effects, which means that parents in these localities send fewer of their kids to schools and could be indicative of a rejection of colonial-like modern education institutions from an intergenerational perspective, although more concrete evidence will need to be provided in the future to further support the hypothesis on the intergenerational transmission of values and norms shaped by the exposure to forced labor (Table 7).

We can see also that the negative effect of military conscription on infrastructure is not specific to schools, as exposure of localities to military service or forced labour also results in lower numbers of health facilities. And such effects on decreased health infrastructure supply are particularly concentrated on the number of pharmacies and maternity centers in the long run (Table 8), which are more likely to be constructed based on community or private initiatives (Chauvet et al., 2015).

5.2.3 Effects by cohorts and gender

Ultimately, with the census data at hand, we also attempt to look at who are the ones getting most affected in terms of human capital accumulation by exposure to forced labor

[50] Currently we are only able to match well the community schools and public schools (around 90%) with our matched village panel (identified from the colonial archives). Quranic schools are rather badly matched (less than 50%) for the moment due to a lot of missing values on the villages they are located in.

[51] This outcome variable could also be understood as the number of students per teacher at the locality level

over the very long run.

Firstly, we might expect that the effects should be the most pronounced among males, as they are the ones (or their forefathers) directly exposed to the conscription regime. However, given that most of our effects would operate in an intergenerational framework, the rejection and mistrust towards embodiment of colonial institutions could be easily transmitted towards both sons and daughters within a given household. As such, a priori, over the long run there's no strong evidence as to why males would be more persistently affected than females. This is indeed what we find in Table 10, where the effects on the incidence of having attended formal primary schooling are significant for both male and female populations. Although in terms of magnitudes, the effect on men appears larger in the early waves of censuses, although such an effect dissipates over time, and the effect on women persists even until 2009. The next natural question to ask is whether the effects are only present for the adult population who themselves experienced the forced labor regime, or if the effects are present for their children, especially on kids during the age range of primary schooling when the censuses are conducted, therefore validating the inter-generational transmission story. On school enrollment outcomes for the children in the age range of 7 to 12, we find that exposure to forced labor has a highly significant and negative impact on girls' schooling in particular. The negative effects are also present for boys, but they are only significant in the 1987 round of census. Furthermore, the effects on both girls and boys become insignificant in the latest round of census in 2009.

Further empirical investigation is needed to better understand if the inter-generational transmission cycle of human capital trap caused by colonial military forced labor has been broken in the most recent years, and if this could be linked to reforms in the education system after democratization in Mali that has seen a rising share of community-initiated schools (Lange & Diarra, 1999) .

5.2.4 Testing mechanisms with Afrobarometer data

Our main hypothesis regarding the lasting effect of exposure to colonial coercive policies on human capital outcomes is that it stems from a rejection of school as an institution

Table 10: Effect of military service and forced labor on education outcomes by cohort and gender

Census 1976

	At least primary schooling						School enrollment					
	School Age and Above		Women		Men		School Age (7-12)		Girls		Boys	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share_Soldiers	0.00	-0.08***	0.00	-0.07***	0.01	-0.10**	0.02***	-0.10***	0.02***	-0.12***	0.01	-0.07
	(0.01)	(0.03)	(0.00)	(0.02)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)	(0.04)	(0.01)	(0.04)
Share_Reservists x District Activation Rate		-0.18***		-0.15***		-0.24***		-0.25***		-0.32***		-0.17*
		(0.07)		(0.05)		(0.09)		(0.07)		(0.08)		(0.09)
Mean DV	0.08	0.08	0.05	0.05	0.11	0.11	0.06	0.06	0.05	0.05	0.07	0.07
St. Dev. DV	0.08	0.08	0.06	0.06	0.10	0.10	0.08	0.08	0.08	0.08	0.10	0.10
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	4276	4276	4273	4273	4271	4271	4237	4237	4215	4215	4215	4215
Adj. R^2	0.08	0.08	0.07	0.07	0.09	0.09	0.06	0.06	0.04	0.05	0.04	0.05

Census 1987

	At least primary schooling						School enrollment					
	School Age and Above		Women		Men		School Age (7-12)		Girls		Boys	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share_Soldiers	0.00	-0.06***	0.00	-0.04***	0.00	-0.08***	-0.00	-0.11***	-0.00	-0.09***	-0.01	-0.12**
	(0.01)	(0.02)	(0.00)	(0.02)	(0.01)	(0.03)	(0.01)	(0.04)	(0.01)	(0.03)	(0.01)	(0.05)
Share_Reservists x District Activation Rate		-0.13***		-0.09***		-0.18***		-0.21***		-0.19***		-0.24**
		(0.05)		(0.04)		(0.06)		(0.08)		(0.07)		(0.10)
Mean DV	0.06	0.06	0.03	0.03	0.08	0.08	0.09	0.09	0.06	0.06	0.11	0.11
St. Dev. DV	0.08	0.08	0.06	0.06	0.10	0.10	0.13	0.13	0.11	0.11	0.16	0.16
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5605	5605	5592	5592	5595	5595
Adj. R^2	0.08	0.08	0.06	0.06	0.09	0.09	0.06	0.06	0.05	0.05	0.06	0.07

Census 1998

	At least primary schooling						School enrollment					
	School Age and Above		Women		Men		School Age (7-12)		Girls		Boys	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share_Soldiers	-0.00	-0.09***	0.00	-0.08***	-0.00	-0.10***	-0.01	-0.15**	-0.00	-0.13**	-0.01	-0.14*
	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.04)	(0.01)	(0.06)	(0.01)	(0.06)	(0.02)	(0.07)
Share_Reservists x District Activation Rate		-0.18***		-0.16***		-0.20***		-0.30**		-0.28**		-0.28*
		(0.06)		(0.05)		(0.07)		(0.13)		(0.12)		(0.15)
Mean DV	0.09	0.09	0.06	0.06	0.12	0.12	0.19	0.19	0.14	0.14	0.23	0.23
St. Dev. DV	0.10	0.10	0.08	0.08	0.13	0.13	0.21	0.21	0.19	0.19	0.23	0.23
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5607	5607	5607	5607	5606	5606	5600	5600	5603	5603
Adj. R^2	0.14	0.15	0.11	0.11	0.16	0.16	0.14	0.14	0.13	0.13	0.14	0.14

Census 2009

	At least primary schooling						School enrollment					
	School Age and Above		Women		Men		School Age (7-12)		Girls		Boys	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Share_Soldiers	-0.02*	-0.13***	-0.01	-0.13***	-0.03**	-0.12**	-0.01	-0.09	-0.01	-0.11	-0.02	-0.08
	(0.01)	(0.05)	(0.01)	(0.04)	(0.01)	(0.05)	(0.02)	(0.08)	(0.02)	(0.08)	(0.02)	(0.08)
Share_Reservists x District Activation Rate		-0.23**		-0.24***		-0.21*		-0.16		-0.20		-0.12
		(0.10)		(0.09)		(0.11)		(0.15)		(0.16)		(0.16)
Mean DV	0.21	0.21	0.17	0.17	0.25	0.25	0.37	0.37	0.34	0.34	0.40	0.40
St. Dev. DV	0.15	0.15	0.13	0.13	0.17	0.17	0.26	0.26	0.25	0.25	0.27	0.27
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	5606	5606	5607	5607	5604	5604	5601	5601	5600	5600
Adj. R^2	0.24	0.24	0.19	0.19	0.28	0.28	0.26	0.26	0.21	0.21	0.30	0.30

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. District fixed effects are colonial fixed effects. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regressions. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table 11: Effect of military service and forced labor on perception of France and traditional leaders

	(1)	(2)	(3)	(4)
	Trust traditional leaders a lot		Barkhane somewhat useful	
Share_Soldiers	-0.06	0.44*	-0.22	-1.00*
	(0.07)	(0.24)	(0.15)	(0.51)
Share_Reservists x District Activation Rate		1.30**		-2.14*
		(0.61)		(1.24)
Constant	0.66***	0.18	0.60***	1.38***
	(0.02)	(0.23)	(0.04)	(0.46)
Mean DV	0.64	0.64		
St. Dev. DV	0.21	0.21		
Dsitrict FE	Yes	Yes	Yes	Yes
Obs	524	524	132	132
Adj. R^2	0.16	0.17	0.37	0.38

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. District fixed effects are colonial fixed effects. Standard errors clustered at the Afrobarometer enumeration area level. Significance levels: *p<0.1; **p<0.05; ***p<0.01

associated with the colonial experience. If such a mechanism is true, we would necessarily observe a larger rejection of the former colonial power in more affected localities. While we currently lack data on the perception of France, we utilize the perception of the usefulness of the French military intervention, Barkhane, as a proxy. Barkhane operated in Mali between 2014 and 2022. Although the ousting of the French army encountered widespread support from the population in 2022, in 2018 during Afrobarometer round 7 data collection, over 53% of Malian respondents reported that the Barkhane operation was at least somewhat useful in the country. As expected, we find that this share was lower in localities affected by colonial military service and even lower in those affected by forced labor. Traditional leaders, on the other hand, tend to be considered as more trustworthy by respondents in localities affected by coercive colonial conscription. While the very small sample size lends caution to interpretation of these results, they would suggest that the colonial experience generated resentment towards the former colonial power but not towards traditional leaders.

6 Conclusion

In this paper, we explore the long-term impacts on local development of a coercive labor regime which relied on military conscripts, drawing evidence from French colonial Mali. From 1927 until the end of the 1940s, the reservists of military conscripts in this former French colony were activated as forced workers for the construction of dams for a gigantic irrigation project and railway renovation.

We combine historic data of handwritten individual military conscript files with four rounds of contemporary Malian censuses to assess the linkage between development and the forced labor system at the contemporary Malian locality level. Through computer-automated processing, we have extracted information on military draft status and localities from the individual files, which we have then matched with the data from the four post-independence Malian censuses.

Relying on the quasi-random allocation of recruited young males into either the regular-soldier or military-reservist categories at the locality level, we find persistent negative impacts military service and forced labour during colonial times. These effects are particularly strong for long-term human capital outcomes. Preliminary evidence suggests that the aggravating human capital outcomes are more likely to be driven by an endogenous demand effect (rejection of modern educational institutions embodied by French colonialism) rather than a supply effect (provision of schools at the locality level), and such effects are transmitted inter-generationally, as children of primary schooling age are less likely to be enrolled in school across different censuses. Although such negative impacts on human capital appear to be gradually fading away in the most recent decades. Further investigation into the mechanisms of potentially worsening inter-generational transmission of human capital is needed, for instance whether education reforms after democratization in Mali may have attenuated the rejection effect of schools in localities exposed to military service and forced labour.

On another note from the supply side of the story, additional research is also called for to better comprehend the institutional interplay between the French colonial administration and the local indigenous chiefs who purportedly played a pivotal role in the provision of public goods, including schools over the very long run.[52] We consider our preliminary results as a first step towards a more comprehensive understanding of the linkage between the forced labor system embedded in military conscription in colonial Mali and local development over the very long run.

[52]The role of chiefs in the conscription system was such that they took part in the conscription system by establishing lists of individuals eligible for recruitment within each village, and made sure the eligible males would present themselves to the drafting centers. Their cooperation or a lack thereof with the French colonial authorities may have impacted their legitimacy in the eyes of the locals in their native villages, which could have resulted in long-run repercussions in their ability to help provide public goods (such as education) pivotal for economic development at the local level.

References

Alesina, A., Hohmann, S., Michalopoulos, S., & Papaioannou, E. (2020). *Religion and educational mobility in africa* (Tech. Rep.). National Bureau of Economic Research.

Alesina, A., Hohmann, S., Michalopoulos, S., & Papaioannou, E. (2021). Intergenerational mobility in africa. *Econometrica*, *89*(1), 1–35.

Archibong, B., & Obikili, N. (2023). *Prison labor: The price of prisons and the lasting effects of incarceration* (Tech. Rep.). National Bureau of Economic Research.

Bernard, C., Mesplé-Somps, S., & Spielvogel, G. (2017). Taille des villes, urbanisation et spécialisations économiques; une analyse sur micro-données exhaustives des 10 000 localités maliennes.

Blouin, A. (2021, 04). Culture and Contracts: The Historical Legacy of Forced Labour. *The Economic Journal*, *132*(641), 89-105. Retrieved from `https://doi.org/10.1093/ej/ueab031` doi: 10.1093/ej/ueab031

Bogosian, C. (2002). Forced labor, resistance and memory: the deuxieme portion in the french soudan, 1926–1950. *Ph.D. Thesis*.

Bogosian, C. (2015). Free to coerce: Forced labour during and after the vichy years in french west africa. *Recentering Africa in World War II*.

Chauvet, L., Gubert, F., Mercier, M., & Mesplé-Somps, S. (2015). Migrants' home town associations and local development in mali. *The Scandinavian Journal of Economics*, *117*(2), 686–722.

Cogneau, D. (2023). *Un empire bon marché: Histoire et économie politique de la colonisation française, xixe-xxie siècle*. Seuil.

Cogneau, D., & Mo, Z. (2023). Rationally coercive rulers? evidence from head tax and blood tax levy in french west africa. *Paris School of Economics, Unpublished Working Paper*.

Delavignette, R. L. (1939). *Les vrais chefs de l'empire*. Gallimard.

Dell, M. (2010). The persistent effects of peru's mining mita. *Econometrica*, *78*(6), 1863–1903.

Dell, M., & Olken, B. A. (2020). The development effects of the extractive colonial economy: The dutch cultivation system in java. *The Review of Economic Studies*, *87*(1), 164–203.

Dell, M., Shen, Z., Zhang, K., & Zhou, J. (2019). Information extraction from text regions with complex tabular structure. *Conference on Neural Information Processing Systems Document Intelligence Workshop*.

Do, Q., Dray, S., Huillery, E., & Keene, J. (2020). Emperors without scepters: Early colonial leaders' personality and civil conflicts.

Dupas, P., Falezan, C., Mabeu, M. C., & Rossi, P. (2023). *Long-run impacts of forced labor migration on fertility behaviors: Evidence from colonial west africa* (Tech. Rep.). National Bureau of Economic Research.

Echenberg, M., & Filipovich, J. (1986). African military labour and the building of the office du niger installations, 1925–1950. *The Journal of African History*, *27*(3), 533–551.

Echenberg, M. J. (1991). Colonial conscripts: The tirailleurs sénégalais in french west africa, 1857-1960.

Echenberg, M. J. (2009). *Les tirailleurs sénégalais en afrique occidentale française, 1857-1960*. KARTHALA Editions.

Fall, B. (1993). *Le travail forcé en afrique occidentale française (1900-1946)* (No. 41). Université Libre de Bruxelles.

Fenske, J. (2013). Does land abundance explain african institutions? *The Economic Journal*, *123*(573), 1363–1390.

Ginio, R. (2017). *The french army and its african soldiers: The years of decolonization*. U of Nebraska Press.

Huillery, E. (2009, April). History matters: The long-term impact of colonial public investments in french west africa. *American Economic Journal: Applied Economics*, *1*(2), 176-215. Retrieved from http://www.aeaweb.org/articles?id=10.1257/app.1.2.176 doi: 10.1257/app.1.2.176

Klein, M. A. (2011). Slaves and soldiers in the western soudan and french west africa. *Canadian Journal of African Studies / Revue Canadienne des Études Africaines*, *45*(3), 565–

587. Retrieved 2023-01-12, from `http://www.jstor.org/stable/23236899`

Lange, M.-F., & Diarra, S. O. (1999). École et démocratie. *Politique africaine*, *N° 76*. doi: 10.3917/polaf.076.0164

Libois, F., Muralidharan, R., Singh, J., & Eynde, O. V. (2020). From muscle drain to brain gain: The long-term effects of gurkha recruitment in nepal.

Lowes, S., & Montero, E. (2021a). Concessions, violence, and indirect rule: evidence from the congo free state. *The Quarterly Journal of Economics*, *136*(4), 2047–2091.

Lowes, S., & Montero, E. (2021b). The legacy of colonial medicine in central africa. *American Economic Review*, *111*(4), 1284–1314.

Magasa, A. (1978). *Papa-commandant a jeté un grand filet devant nous: les exploités des rives du niger 1900-1962*. FeniXX.

Mann, G. (2006). *Native sons*. Duke University Press.

Mesplé-Somps, S., Selod, H., Spielvogel, G., & Blankespoor, B. (2018). Urbanisation et croissance dans les villes du mali.. doi: 10.4000/books.irdeditions.21242

Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of Mistrust in Africa. *American Economic Review*. doi: 10.1257/aer.101.7.3221

Pruett, L. (2024). Resisting the blood tax: Coercive capacity, railroads and draft evasion in colonial west africa. *The Journal of Politics*.

Renault, F. (1972). *L'abolition de l'esclavage au sénégal,*. Société française d'histoire d'outre-mer, P. Geuthner.

Ronconi, J. P., & Ramos-Toro, D. (2022). Nation-building through military service.

Salem, A. (2022). *The negative impacts of colonization on the local population: Evidence from morocco* (Tech. Rep.). Working Paper.

Salem, A., & Seck, A. A. (2022). En route: The french colonial army, emigration, and development in morocco.

Tiquet, R. (2019). *Travail forcé et mobilisation de la main-d'oeuvre au sénégal: années 1920-1960*. Presses universitaires de Rennes.

Tiquet, R. (2020). Education through labor: From the deuxième portion du contingent to the youth civic service in west africa (senegal/mali, 1920s–1960s). *Education and Development in Colonial and Postcolonial Africa: Policies, Paradigms, and Entanglements,* *1890s–1980s*, 83–105.

Vanden Eynde, O. (2016). Military service and human capital accumulation: Evidence from colonial punjab. *Journal of Human Resources*, *51*(4), 1003–1035.

Van Waijenburg, M. (2018). Financing the african colonial state: The revenue imperative and forced labor. *The Journal of Economic History*, *78*(1), 40–80.

Zuccarelli, F. (1962). Le régime des engagés à temps au sénégal (1817-1848). *Cahiers d'études africaines*, 420–461.

A Appendix

A.1 Functioning of Conscription System

Figure A1: Top-down Military Conscription Procedures: Military Quota Allocation

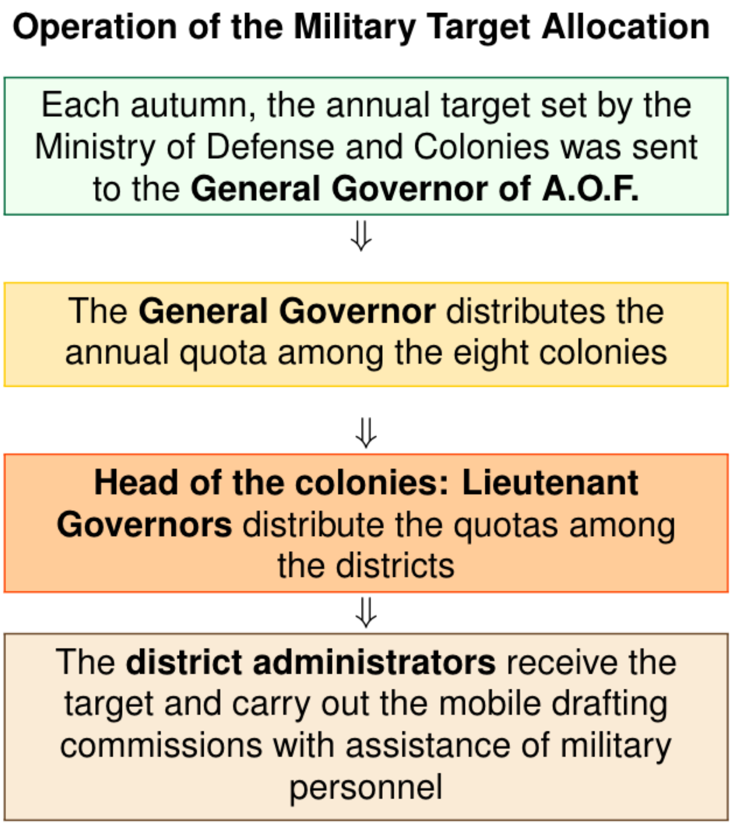
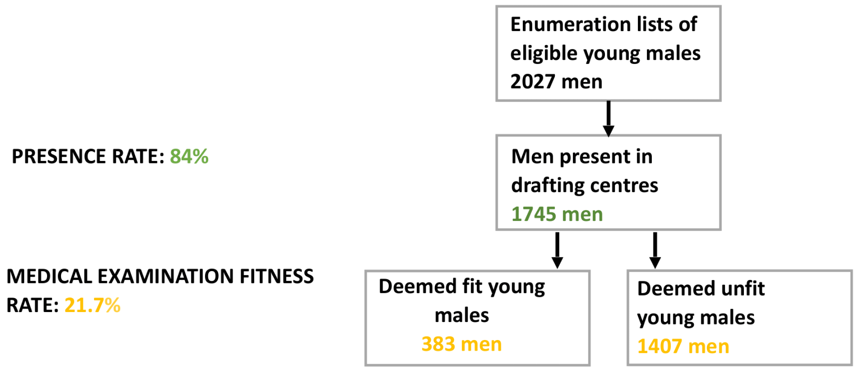
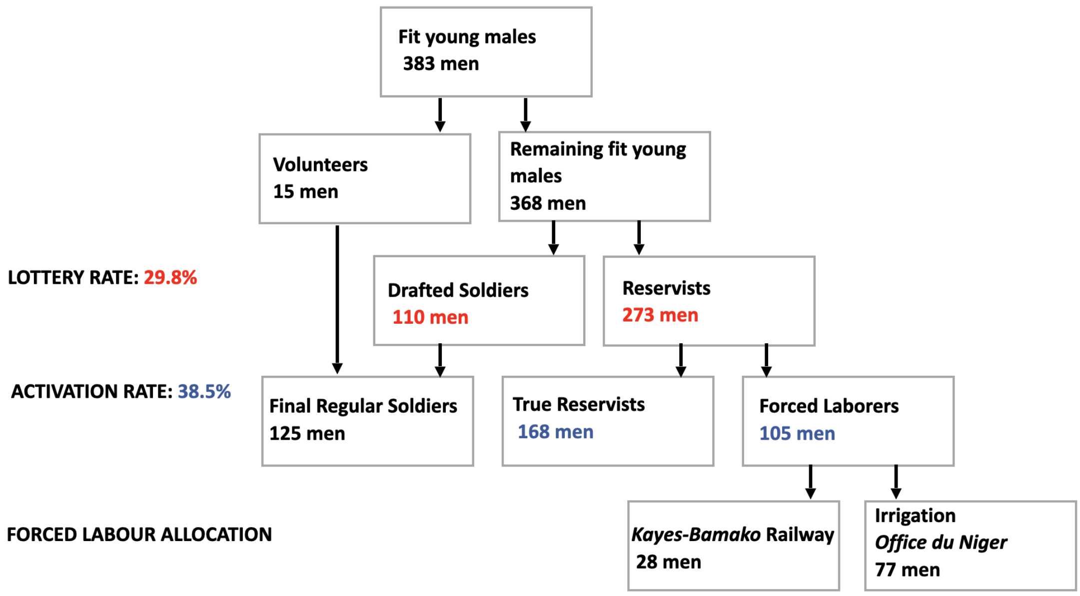


Figure A2: Simulated Mobile Drafting Commission Procedure with Forced Labor Activation

(a) Procedure (1)



(b) Procedure (2)



Notes: The whole drafting procedure described here took place at the colonial district level on an annual basis.

A.2 Additional Tables and Figures

Table B1: Summary of Key Military Procedural Variables with ANS District-level Data

VARIABLES	(1) N	(2) Mean	(3) S.D.	(4) Min	(5) Max
Raw Military Quota	360	110.5	83.70	0	400
Re-incorporated Absentees	407	6.351	17.13	0	159
Military Quota	391	107.3	81.75	0	400
Enumerated	407	2,027	1,489	0	7,902
Present	407	1,745	1,362	0	7,046
Absentees	378	303.8	376.5	0	3,110
Fit	407	383.5	364.7	0	2,875
Unfit	392	1,407	1,131	0	6,312
Incorporated (Drafted) Soldiers	407	116.6	88.70	0	500
First-Portion Regular Soldiers	407	110.4	86.81	0	500
Volunteer Soldiers	407	15.18	30.26	0	279
Appelés	407	95.18	82.27	0	493
Second-Portion Reservists	407	273.1	328.7	0	2,795
Reservist Forced Laborer	328	104.6	114.2	0	734
Reservist Forced Laborer - ***Office du Niger***	297	62.36	92.92	0	620
Reservist Forced Laborer - ***Kayes-Bamako Railway***	297	22.79	62.00	0	567
Absenteeism Rate	361	0.160	0.145	0.000299	0.792
Present Rate	361	0.840	0.145	0.208	1.000
Volunteering Rate	356	0.163	0.219	0	1
Deemed Fit Rate	361	0.217	0.116	0.0514	0.817
Lottery Rate	361	0.669	0.195	0	1

Table B2: Impact of Second Portion Forced Labor on Log Volunteering Outcomes

VARIABLES	(1) Volunteer Rate$_t$	(2) Volunteer Rate$_t$	(3) Volunteer Rate$_t$
Activation Rate$_t$	-0.0067	-0.0055	0.0114
	(0.0303)	(0.0313)	(0.0335)
Activation Rate$_{t-1}$	0.0156	0.0162	0.0235
	(0.0323)	(0.0346)	(0.0387)
Activation Rate$_{t-2}$	0.1084**	0.1077**	0.1000**
	(0.0433)	(0.0430)	(0.0429)
Present Rate$_t$		0.1324	0.0627
		(0.0802)	(0.0792)
Fit Rate$_t$		-0.0787	-0.2972**
		(0.0824)	(0.1385)
Lottery Rate$_t$			0.2446**
			(0.0897)
Observations	323	323	323
R-squared	0.536	0.541	0.558
District FE	YES	YES	YES
Year FE	YES	YES	YES

Notes: Panel regressions at the colonial district-year level with ANS data. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Figure B1: Example of Individual Soldier Files from Military Archives at Pau (CAPM)

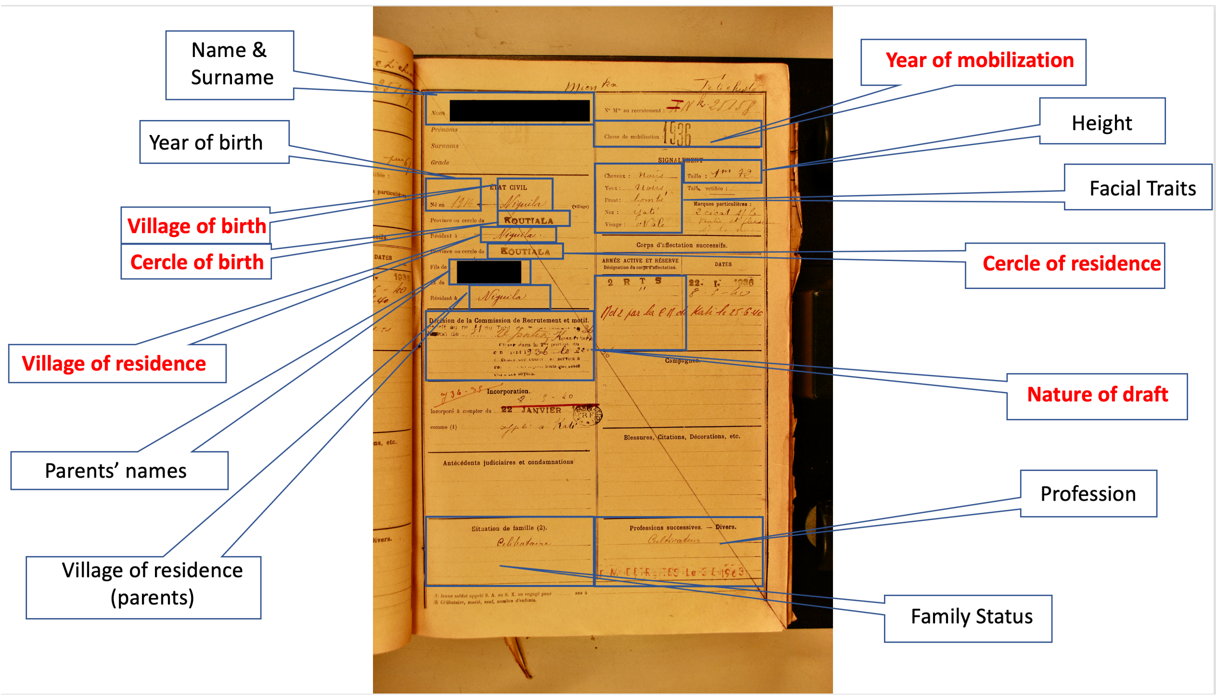


Table B3: Impact of Second Portion Forced Labor on Volunteering Outcomes with PPML Model

VARIABLES	(1) Volunteers	(2) Volunteers	(3) Volunteers
ihs(Forced Laborer)t	0.1385*	0.1315*	0.1102
	(0.0741)	(0.0717)	(0.0794)
ihs(Forced Laborer)t-1	0.0635*	0.0520	0.0586
	(0.0380)	(0.0381)	(0.0403)
ihs(Forced Laborer)t-2	0.0756**	0.0555	0.0566
	(0.0375)	(0.0389)	(0.0382)
ihs(Soldiers Drafted)t		0.3898	0.3785
		(0.2761)	(0.2765)
ihs(Military Quota)t		0.5290	0.5540
		(0.3769)	(0.3942)
ihs(Total Reservists)t			0.0977
			(0.0851)
Observations	320	319	319
District FE	YES	YES	YES
Year FE	YES	YES	YES

Notes: Panel regressions with Pseudo-Poisson Maximum Likelihood at the colonial district-year level with ANS data. We take the inverse hyperbolic form (ihs) of the explanatory variables. Robust standard errors in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table B4: Temporal correlation in conscription outcomes

	(1) $Share\_Reservist_t$	(2) $Share\_Reservist_t$	(3) $Tot\_Conscripts_t$	(4) $Tot\_Conscripts_t$	(5) $Volunteers_t$	(6) $Volunteers_t$
$Share\_Reservist_{t-1}$	0.0341***	0.0083				
	(0.0100)	(0.0097)				
$Tot\_Conscripts_{t-1}$			0.7186***	0.7187***		
			(0.0218)	(0.0215)		
$Volunteers_{t-1}$					0.2326***	0.2291***
					(0.0342)	(0.0345)
Mean DV	0.718	0.718	6.941	6.941	0.262	0.262
Colonial District FE	No	Yes	No	Yes	No	Yes
Obs	14243	14243	14341	14341	14341	14341
Adj. R^2	0.001	0.041	0.507	0.522	0.034	0.045

Notes: Locality-year level data grouped in four periods of equal number of conscripts (1927-1931, 1932-1938, 1939-1943,1944-1950. Only last three periods included in sample as lagged values are not available for the first period.

Table B5: Spatial correlation in conscription outcomes

	(1) Share Soldiers	(2) Share Soldiers	(3) Activation Rate	(4) Activation Rate	(5) Share Volunteers	(6) Share Volunteers
Neighbors : Share Soldiers	0.14***	-0.00				
	(0.02)	(0.02)				
Neighbors : Activation Rate			0.21***	0.02		
			(0.02)	(0.02)		
Neighbors : Share Volunteers					0.13***	0.07***
					(0.02)	(0.02)
Constant	0.22***	0.26***	0.66***	0.82***	0.03***	0.04***
	(0.00)	(0.00)	(0.02)	(0.02)	(0.00)	(0.00)
Mean DV	0.26	0.26	0.83	0.83	0.04	0.04
St. Dev. DV	0.18	0.18	0.29	0.29	0.08	0.08
Year FE	No	Yes	No	Yes	No	Yes
Obs	5573	5573	3168	3167	5594	5594
Adj. R^2	0.01	0.10	0.04	0.17	0.01	0.05

Notes: Locality level data is available for all years excepting the distribution of reservists between true reservists and forced workers (i.e. the activation rate) available for 1941, 1942, 1944 and 1945 only. Neighbors include all localities within a radius of 5km. Standard errors clustered at the locality level. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Figure B2: Example of District-level Conscription Tables for Colonial Mali taken from Archives Nationales du Sénégal (ANS), Dakar

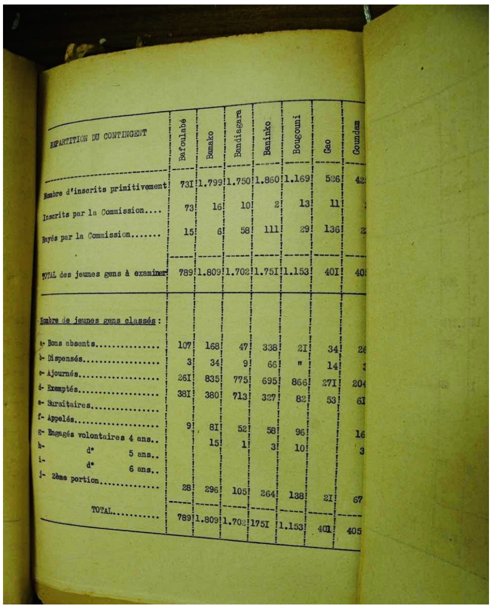


Table B6: Locality level determinants of recruitment of reservists and soldiers - ethnic composition

	Activation		Share_Reservist	
	(1)	(2)	(3)	(4)
Malinke	0.06**	0.06**	−0.03***	−0.01
	(0.03)	(0.03)	(0.01)	(0.01)
Peul/foulfoulbe	0.06**	0.05**	−0.000	0.003
	(0.02)	(0.03)	(0.01)	(0.01)
Sonrai/djerma	0.01	−0.01	−0.05***	−0.01
	(0.04)	(0.07)	(0.01)	(0.01)
Maraka/soninke	−0.07**	−0.01	−0.05***	−0.01
	(0.03)	(0.04)	(0.01)	(0.01)
Kassonke	0.12***	0.04	−0.05***	−0.003
	(0.04)	(0.05)	(0.02)	(0.02)
Senoufo	0.08**	0.01	0.06***	0.03***
	(0.03)	(0.05)	(0.01)	(0.01)
Dogon	0.20***	0.18***	−0.01*	0.002
	(0.02)	(0.03)	(0.01)	(0.01)
Bobo	0.35***	0.07***	−0.12***	−0.02*
	(0.02)	(0.02)	(0.01)	(0.01)
Minianka	−0.01	−0.10*	0.04***	0.03**
	(0.05)	(0.05)	(0.01)	(0.01)
other	0.05	0.02	0.02	0.01
	(0.05)	(0.05)	(0.01)	(0.01)
Mean DV	0.73		0.72	
Year FE	Yes	Yes	Yes	Yes
District-year FE	No	Yes	No	Yes
Obs	4,243	4,243	55,496	55,496
Adj. R^2	0.06	0.24	0.10	0.21

Notes: The ethnicity variables are dummies that take on value one if the largest (most populous) ethnic group in a given village is the said ethnic group. Thee reference group is Bamabara (the largest ethnicity in Mali). We include colonial district by year fixed effects. Locality level data is available for all years excepting the distribution of reservists between true reservists and forced workers (i.e. the activation rate) available for 1941 and 1942 only. Standard errors clustered at the locality level. Ethnic composition of localities is measured with the 2009 population census. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table B7: Summary statistics localities - 1976 census

Statistic	N	Mean	St. Dev.	Min	Max
		Archives variables			
Total absentees	5,621	0.071	0.321	0	5
Total volunteers	5,621	0.827	2.107	0	64
Total drafted soldiers	5,621	5.729	7.423	0	195
Total soldiers	5,621	6.557	8.786	0	231
Share of Soldiers	5,621	0.289	0.181	0.000	1.000
Total true reservists	5,621	0.433	1.143	0	28
Total forced-laborers	5,621	1.920	2.808	0	43
Total reservists n.c.e.	5,621	13.939	16.085	0	181
Total reservists	5,621	16.292	18.455	0	210
Share reservists	5,621	0.703	0.185	0.000	1.000
		Censuses variables			
Number of Hhs	5,620	130.748	173.288	1	3,167
Total residents	5,621	662.091	894.148	3	16,082
Female residents	5,621	340.678	459.636	2	8,014
Male residents	5,621	321.413	435.895	1	8,068
Fertility	4,282	0.437	0.139	0.000	4.000
No education	4,287	0.923	0.078	0.000	1.000
Primary school	4,287	0.077	0.078	0.000	1.000
Secondary school	4,287	0.001	0.004	0.000	0.139
University	4,287	0.0001	0.001	0.000	0.044

Notes: Observations differ due to missingness.

Table B8: Summary statistics localities - 1987 census

Statistic	N	Mean	St. Dev.	Min	Max
		Archives variables			
Total absentees	5,621	0.071	0.321	0	5
Total volunteers	5,621	0.827	2.107	0	64
Total drafted soldiers	5,621	5.729	7.423	0	195
Total soldiers	5,621	6.557	8.786	0	231
Share of Soldiers	5,621	0.289	0.181	0.000	1.000
Total true reservists	5,621	0.433	1.143	0	28
Total forced-laborers	5,621	1.920	2.808	0	43
Total reservists n.c.e.	5,621	13.939	16.085	0	181
Total reservists	5,621	16.292	18.455	0	210
Share reservists	5,621	0.703	0.185	0.000	1.000
		Censuses variables			
Number of Hhs	5,621	138.956	198.597	1	3,071
Total residents	5,621	734.582	1,103.911	13	18,379
Female residents	5,621	383.473	566.575	6	9,272
Male residents	5,621	351.110	539.035	5	9,107
Fertility	5,621	0.885	0.217	0.000	2.617
No education	5,621	0.940	0.078	0.399	1.000
No formal education	5,621	0.943	0.077	0.400	1.000
Primary school	5,621	0.057	0.077	0.000	0.600
Secondary school	5,621	0.002	0.006	0.000	0.129
University	5,621	0.0002	0.001	0.000	0.038
Average welfare index	5,621	−0.153	0.431	−2.191	2.521
Absolute within locality gini	5,620	0.165	0.116	0.000	0.767
Inter-quartile range	5,621	0.340	0.391	0.000	2.600

Notes: Observations differ due to missingness.

Table B9: Summary statistics localities - 1998 census

Statistic	N	Mean	St. Dev.	Min	Max
		Archives variables			
Total absentees	5,621	0.071	0.321	0	5
Total volunteers	5,621	0.827	2.107	0	64
Total drafted soldiers	5,621	5.729	7.423	0	195
Total soldiers	5,621	6.557	8.786	0	231
Share of Soldiers	5,621	0.289	0.181	0.000	1.000
Total true reservists	5,621	0.433	1.143	0	28
Total forced-laborers	5,621	1.920	2.808	0	43
Total reservists n.c.e.	5,621	13.939	16.085	0	181
Total reservists	5,621	16.292	18.455	0	210
Share reservists	5,621	0.703	0.185	0.000	1.000
		Censuses variables			
Number of Hhs	5,621	159.733	246.804	1	4,839
Total residents	5,621	934.304	1,464.645	3	28,114
Female residents	5,621	474.543	732.871	1	13,993
Male residents	5,621	459.761	732.903	2	14,135
Fertility	5,621	0.817	0.182	0.000	2.217
No education	5,621	0.866	0.130	0.074	1.000
No formal education	5,621	0.906	0.105	0.345	1.000
Primary school	5,621	0.090	0.100	0.000	0.591
Secondary school	5,621	0.002	0.005	0.000	0.081
University	5,621	0.002	0.008	0.000	0.398
Average welfare index	5,621	−0.306	0.259	−1.130	2.420
Absolute within locality gini	5,620	0.107	0.103	0.000	1.008
Inter-quartile range	5,621	0.173	0.313	0.000	4.267

Notes: Observations differ due to missingness.

Table B10: Summary statistics localities - 2009 census

Statistic	N	Mean	St. Dev.	Min	Max
			Archives variables		
Total absentees	5,621	0.071	0.321	0	5
Total volunteers	5,621	0.827	2.107	0	64
Total drafted soldiers	5,621	5.729	7.423	0	195
Total soldiers	5,621	6.557	8.786	0	231
Share of Soldiers	5,621	0.289	0.181	0.000	1.000
Total true reservists	5,621	0.433	1.143	0	28
Total forced-laborers	5,621	1.920	2.808	0	43
Total reservists n.c.e.	5,621	13.939	16.085	0	181
Total reservists	5,621	16.292	18.455	0	210
Share reservists	5,621	0.703	0.185	0.000	1.000
			Censuses variables		
Number of Hhs	5,621	219.657	407.747	1	14,956
Total residents	5,621	1,352.355	2,552.638	6	96,173
Female residents	5,621	685.834	1,277.042	4	48,119
Male residents	5,621	666.521	1,276.607	2	48,054
Fertility	5,619	0.940	0.270	0.000	11.700
No education	5,621	0.767	0.162	0.000	1.000
No formal education	5,621	0.773	0.160	0.000	1.000
Primary school	5,621	0.212	0.146	0.000	1.000
Secondary school	5,621	0.011	0.019	0.000	0.530
University	5,621	0.004	0.009	0.000	0.196
Average welfare index	5,621	−0.410	0.316	−1.339	2.162
Absolute within locality gini	5,620	0.142	0.094	0.00000	0.656
Inter-quartile range	5,621	0.288	0.288	0.000	2.252

Notes: Observations differ due to missingness.

Figure B3: Histograms of locality-level share of soldiers by districts in colonial Mali

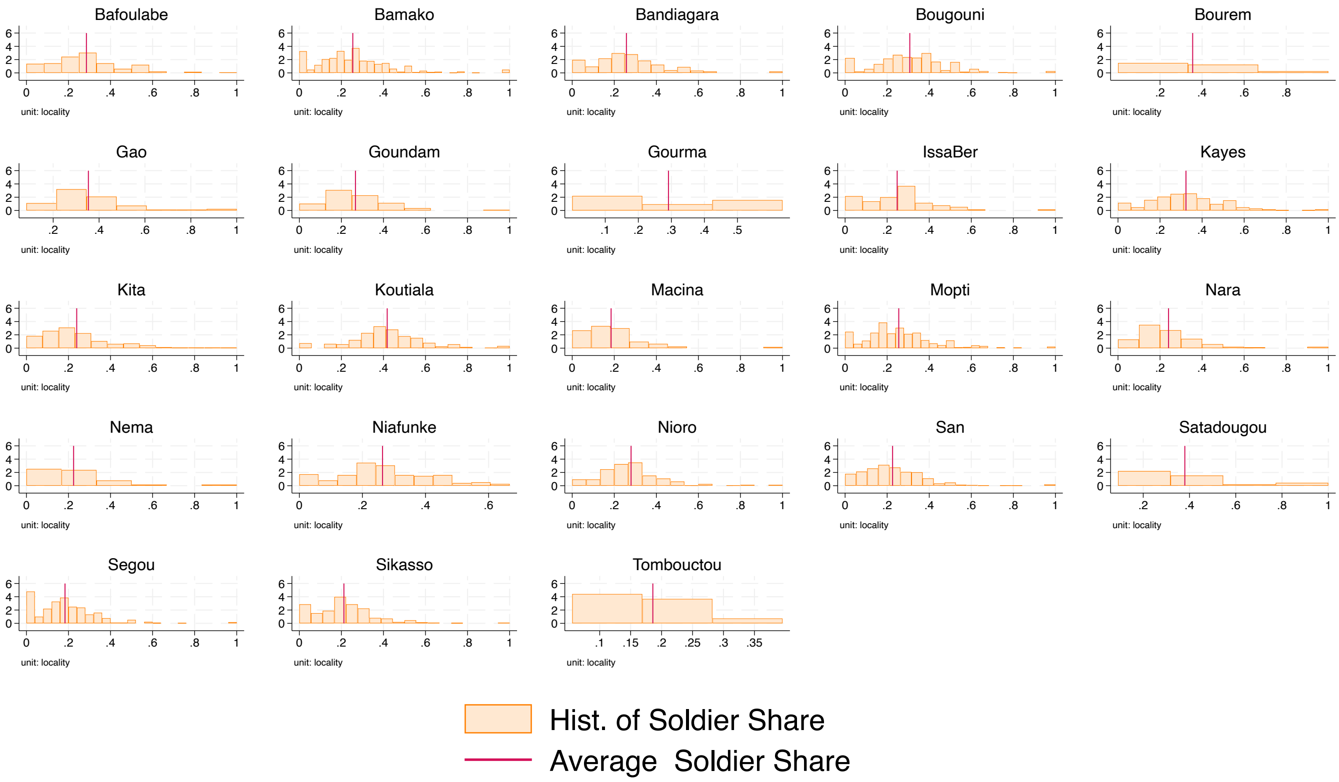
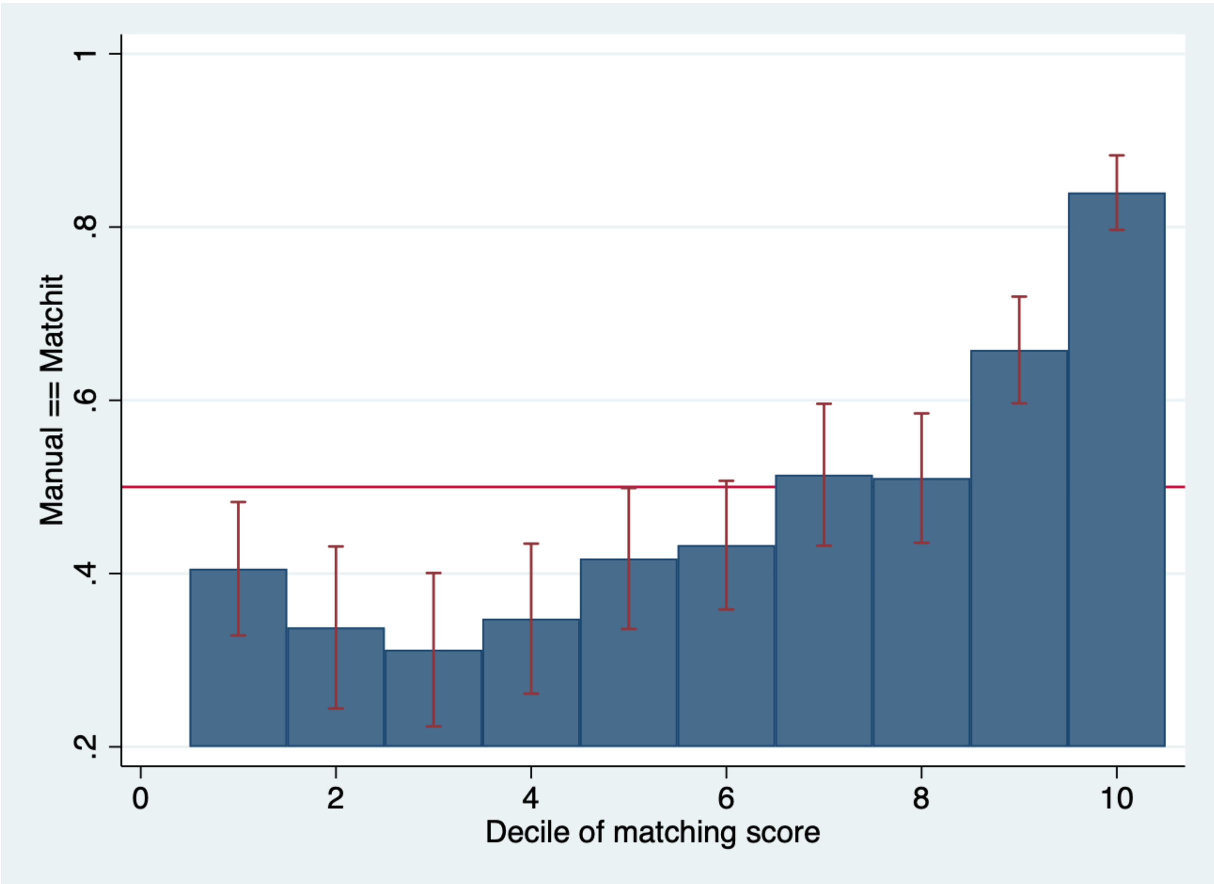


Figure B4: Bar graphs of average rate of agreement between manual choice of village and matchit algorithm. The average score in the 7th decile is 0.68 *Notes:* The assessment is made with about 1200 village entries.



A.3 CAPM Data Processing

A.4 Information Extraction from Individual soldiers files

There has been an increasing interest in document analysis in recent years, with significant progress in automatic layout segmentation of document images into text and non-text regions although efforts to parse the information in text region have been limited (Dell, Shen, Zhang, & Zhou, 2019). As such a contribution of this data collection effort is to exploit a historical database entirely digitized with machine learning algorithms. We worked with Teklia (an automatic document processing company with experience with historical documents), which developed a deep learning model to extract information from the photos

Description of the model

The model developed for this task is a page-level key-value extraction model. This model is trained directly on the image of the double page and is trained to predict the information on either the left or the right page, based on an instruction given to the model during inference.

Training of the model

The model was trained in 2 steps. First, the model was trained on the 3756 annotated single pages with a first set of information to extract. The model was then fine-tuned on the 752 double pages with their left and/or right annotations.

Evaluation

The evaluation was conducted during the training phase on a set of 97 pages not used for training. The necessary condition for this evaluation to be applicable to the population is that the sampling was random. The results of the prediction model for each entity type, on the test set, are given in Table C1.

Table C1: Results of the prediction model for AI entity detection results - CAPM Data

Entity type	Precision	Recall	F1	Support
À compter du	0.938	0.958	0.948	143
Antécédants Judiciaires et condamnations	0.973	0.993	0.983	144
Classe de mobilisation	1.0	1.0	1.0	147
Comme	0.894	0.929	0.911	253
Corps d'affectations sucessifs	0.721	0.736	0.729	144
Decision de la commission de recrutement et motif	0.923	0.876	0.899	177
Et de (mère)	0.822	0.844	0.833	295
Fils de (père)	0.867	0.867	0.867	255
Grade	0.7	0.683	0.691	41
Né à (village)	0.839	0.839	0.839	168
Né en	0.993	0.993	0.993	145
Nom	0.948	0.935	0.941	292
Pays	1.0	1.0	1.0	147
Prénoms	0.81	0.895	0.85	19
Professions successives - Divers	0.906	0.893	0.899	140
Province ou cercle de	0.966	0.96	0.963	149
Province ou cercle de.1	0.973	0.953	0.963	149
Résidant à	0.796	0.792	0.794	168
Résidant à.1	0.802	0.812	0.807	165
Situation de famille	0.863	0.936	0.898	141
Taille	0.833	0.848	0.841	112
All	0.891	0.898	0.895	3394

Notes: Precision/Recall/F1/Support details at https://scikitlearn.org/stable/modules/generated/sklearn.metrics.precision_recall_fscore_support.html

A.5 Matching of village between CAPM data and census

The individual conscript files contain residency and birth location at the village and the district level. The first step consists in cleaning the districts in each of the field transcribed automatically. This is done for 98.7% of the files.[53]

Then, we match village names within a district in the CAPM data to village names within the corresponding district in the 1976 census. Throughout the colonial period, the boundaries of the districts changed, so we make a grouping of districts. This results in about 200 villages with the same names in different arrondissement (a sub-administrative unit within a district) to appear as duplicates. For these cases, we keep the village with the largest population. Table C2 below shows the number of village entries by grouping of districts of residence and the list of villages in the census 1976.

Table C2: Number of village entries by group of districts of residence in the CAPM data and number of villages in the census 1976 data to which they are matched to.

	CAPM data		1976 census	
Districts grouping	Frequency	Percent	Frequency	Percent
Bamako, Nara, Nioro	14,792	21.86	2,297	22.06
Bandiagara, Mopti	11,357	16.78	1,659	15.93
Bougouni	4,730	6.99	834	8.01
Gao, Gourma, Hombori, Bourem	1,979	2.92	472	4.53
Goundam	2,521	3.72	245	2.35
Kita	1,830	2.70	299	2.87
Koutiala	3,818	5.64	540	5.1
Niafunke, IssaBer	2,749	4.06	448	4.30
San	5,178	7.65	730	7.01
Satadougou, Bafoulabe, Kayes	5,585	8.25	768	7.38
Segou, Macina	8,644	12.77	1,408	13.52
Sikasso	3,689	5.45	614	5.90
Tombouctou	806	1.19	99	0.95
Total	67,678	100.0	10,413	100.0

Notes: An observation in the CAPM data is a unique village entry while it is a locality in the 1976 census.

Based on this grouping in both the CAPM and the census, we conduct a fuzzy matching.

[53]Among the files without district information, 2,368 had either village of residence or village of birth noted on them, which are not part of the matching for now as they only represent 0.89% of the files at disposal. We will review these cases time permitting.

A direct merging based on names is unlikely to succeed as village names are transcribed by colonial administration clerks. Hence, there are misspellings and non standardized transcription of villages names from local languages into French.

We use the STATA command *matchit* to perform the matches of villages. We use the default algorithm Ngrams which decomposes the strings into sub-strings to make the comparisons. We prefer this method to the one based on phonetics which are more adapted to the English language. We consider the matches with a similarity score above 0.8 as valid matches, which we consider to be a conservative threshold based on a qualitative assessment of the matches.

Table C3 provides a summary of the matching results. The CAPM database contains approximately 70,000 unique village entries, which we attempted to match with the list of around 10,000 villages from the 1976 census. We did not expect to achieve a perfect match between the two datasets, as some villages may have disappeared or changed names over time. Based on a qualitative assessment of the matches, we consider a match viable if the matching score is between 0.8 and 1. To evaluate the quality of our matching process further, we conducted additional tests. By randomly selecting 1,000 villages (stratified by colonial districts) and manually coding them, we compared our results with those produced by the algorithm. As shown in Figure B4, the rate of disagreement between manual coding and the algorithm's selections falls below 0.5 when the score is under the 70^{th} percentile, where the average score is 0.69. This indicates that using a 0.8 threshold for our algorithm is a conservative approach. The matching exercise yielded valid matches for 25 percent of village entries (out of the approximately 70,000 total unique village entries, when combining villages of residence and birth) on 84,001 individual files from the CAPM database, representing 46.5% of the entire CAPM individual sample.[54]

In addition, in some cases, entries from the archival records with the different handwritten village names could be matched to the same village from the 1976 census, as such the

[54]There is a considerable margin of improvement in the matches especially for the range of matches with matching scores between 0.5 and 0.8. We conducted a manual coding of the most frequent village entries below 0.8 and have also included them in the sample and are assessing ways to make more improvements.

Table C3: Fuzzy matching summary - Matching of names of villages in the CAPM data to the list of the 1976 census

	Locality (resid)		Locality (born)		Individual	
Score	N	%	N	%	N	%
Missing locality					2,368	1.31
No match	11,276	19.39	13,208	21.05	24,213	13.40
Low matching score ($0.5 \leq s < 0.8$)	32,337	55.59	33,821	53.89	70,046	38.78
Matched ($0.8 \leq s \leq 1$)	14,553	25.02	15,728	25.06	84,001	46.50
Total	58,166		62,757		180,628	

Notes: The matching is done by matching the name of the village for each soldier file to a list of villages from the 1976 census. The matching is done by grouping soldier files and villages entries from the census by group of colonial districts.

matching score for one entry could fall below the threshold we set for being considered as a valid match (0.8), but for the other it could be above the valid matching score threshold. In these scenarios we treat these lower-score archival village entries as matched village entries as well.[55] As such, the number of files in the matched sample reported below in Table 2 represents about 80% of the total CAPM database, as opposed to the very restrictive 46.5% matching rate reported earlier.

A.6 Balance tests matching

Table C4 compares the matching rate across colonial districts and the correlation with district matching rate and activation rate in the full sample and excluding the North districts.

A.7 Comparison between ANS data and CAPM data

Figure C1 displays the annual colony-level total numbers of military reservists with the district-level ANS data and the individual-level CAPM data. Rather strikingly, the two series coincide with each other very closely over time, with a small gap towards the end of the 1940s. In terms of the regular soldier numbers in the same graph, we could see that

[55] For instance, both the archival village entries "Goury" and "Gouriumera" are matched to the 1976-census village name "Gouroumera", with the matching score between "Goury" and "Gouroumera" standing at 0.6, while the score between "Gouriumera" and "Gouroumera" at 0.9. Under circumstances of such, we keep the archive village entry "Goury" and treat it as the 1976-census village of "Gouroumera" as well.

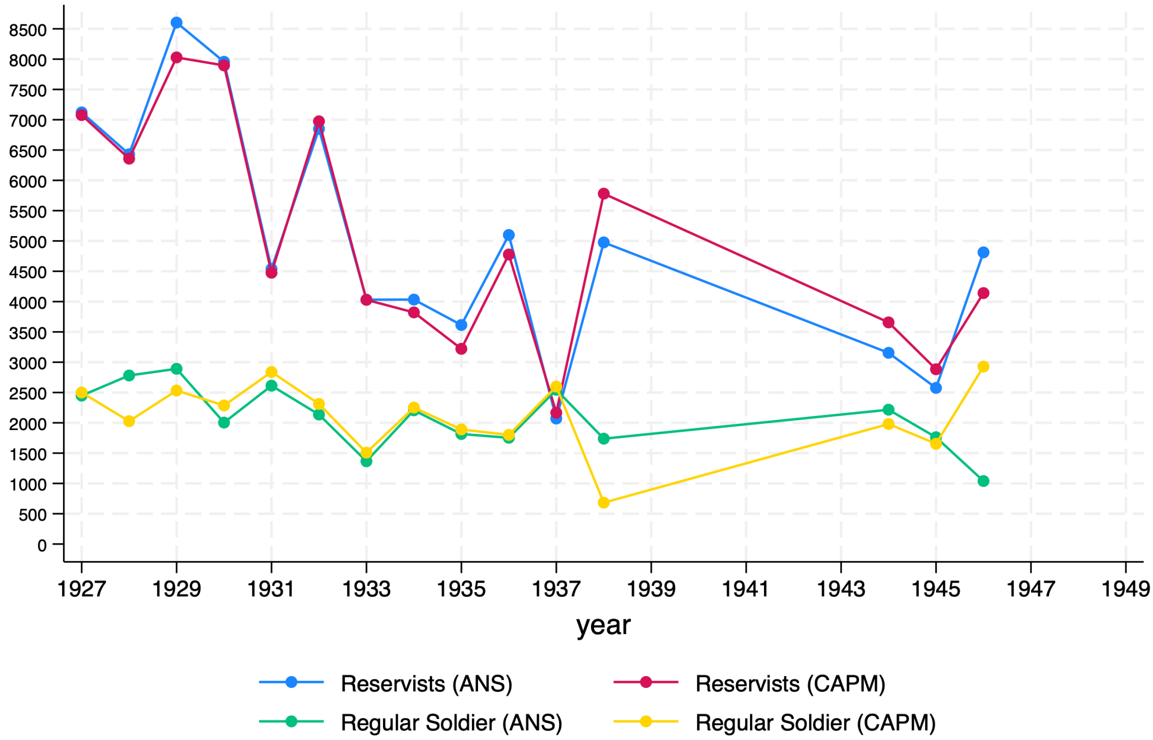
Table C4: District-level determinants of matching rate

	(1)	(2)	(3)
	Archive obs. matched (dummy)		
	All Districts		Excl. North Districts
District Activation Rate		0.08***	0.00
		(0.01)	(0.01)
Bafoulabe	0.00		
	(.)		
Bamako	-0.04***		
	(0.01)		
Bandiagara	-0.08***		
	(0.01)		
Bougouni	-0.00		
	(0.01)		
Bourem	-0.17***		
	(0.02)		
Gao	-0.13***		
	(0.01)		
Goundam	-0.26***		
	(0.01)		
Gourma	-0.25***		
	(0.02)		
Issa-Ber	-0.04***		
	(0.01)		
Kayes	-0.04***		
	(0.01)		
Kita	-0.03***		
	(0.01)		
Koutiala	-0.09***		
	(0.01)		
Macina	-0.02**		
	(0.01)		
Mopti	-0.05***		
	(0.01)		
Nara	-0.04***		
	(0.01)		
Nema	-0.00		
	(0.01)		
Niafunke	-0.05***		
	(0.01)		
Nioro	0.01		
	(0.01)		
San	-0.09***		
	(0.01)		
Satadougou	-0.05***		
	(0.02)		
Segou	-0.04***		
	(0.01)		
Sikasso	0.01		
	(0.01)		
Tombouctou	-0.56***		
	(0.01)		
Constant	0.83***	0.73***	0.81***
	(0.01)	(0.01)	(0.01)
Mean DV	0.78	0.78	0.80
St. Dev. DV	0.41	0.41	0.40
Year FE	Yes		
Obs	179039	150644	140321
Adj. R^2	0.03	0.00	0.00

Notes: An observation represents a conscript file. Activation rate is defined at the colonial district- year level with ANS data. Robust standard errors in parentheses.

the correlation is slightly weaker compared to the reservist correlation, yet for a sub-set of years (1928-1937) the year-to-year variations follow each other very closely as well.[56]

Figure C1: Correlations of Conscripts Counts between ANS data and CAPM data



Notes: This graph shows the annual total of military reservists (both regular soldier and reservist recruits) for the entire colony of Mali, for the years that are non-missing in both ANS data and CAPM data. It concerns the entire sample for both datasets.

A.8 Monte-Carlo simulation of the drafting procedure

We have conduct a Monte-Carlo simulation to emulate the actual conscript drafting procedure back in colonial Mali. We start with the dataset with matched villages only, and keep individuals who are either regular soldier recruits (first-portion) or reservists (second-portion), while not including the volunteers. Specifically, for each year:

[56] ANS data is missing for the period 1939-1943, which more or less coincides with the World War Two period. The enlarging gap between the two data sources in 1938 could be due to the initiation of war-time mobilisation, which will be investigated further in the future.

1. At the colonial district level, we compute the share of reservists (out of the sum of regular soldiers and reservists) using the military draft variable we constructed based on the individual archive dataset.[57]

2. We randomly draw a number between 0 and 1 from a uniform distribution for each conscript in the complete sample of our dataset, for each colonial district in a given year. If this number falls below the actual share of reservists at the colonial district level, the individual conscript is assigned the status of being drafted as a military reservist (as opposed to being drafted as a regular soldier).[58]

3. We repeat this procedure independently for 200 times in total.

4. For each simulation, we calculate the share of individuals being assigned reservist status at the locality level, as well as the average share of the entire simulated distribution.

5. We keep only the post-colonial localities that are well matched, and evaluate the statistical first and second moments differences between the simulated distribution and the actual observed distribution accordingly.[59]

Figure C2 presents a histogram of the share of reservists in the archival dataset across localities, accompanied by a kernel density curve (in blue) of the observed reservist-share distribution. In addition, this figure also plots 50 kernel densities for the simulated distributions of the reservist shares. Overall, we could observe that the observed distribution and the simulation distributions overlay quite well with one another. Although there

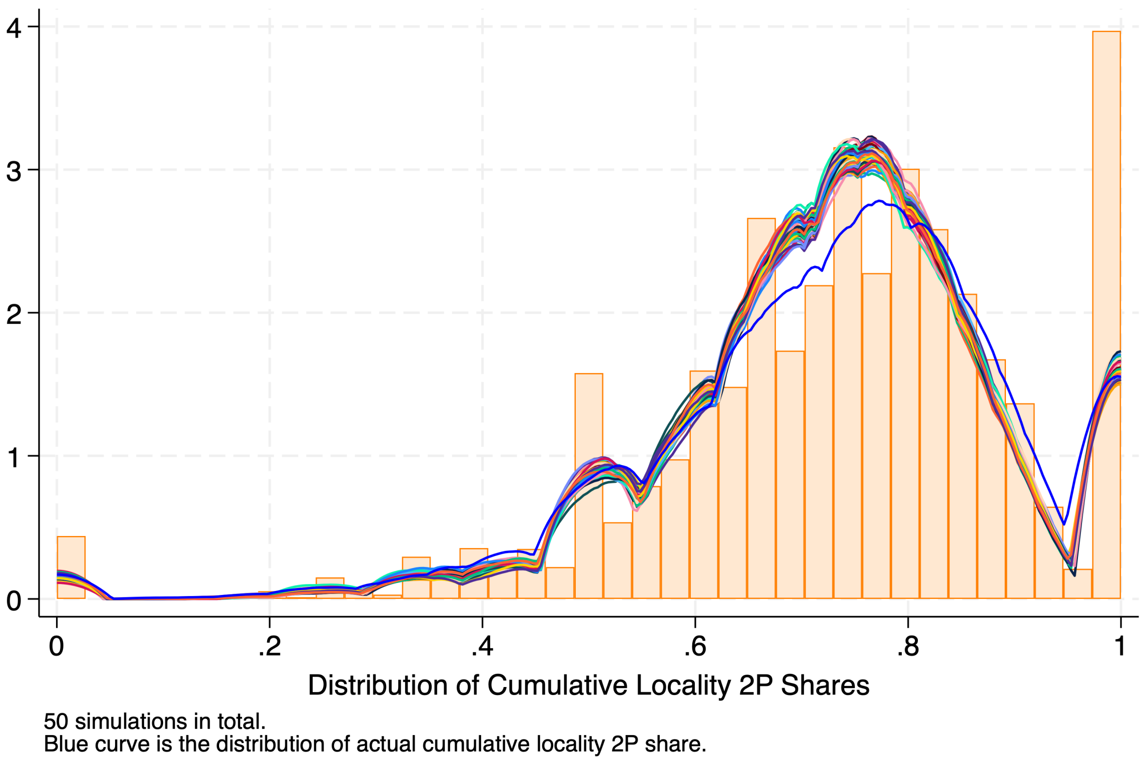
[57]This rate, which was essentially the colony-district-level lottery rate, was set by colonial district administrators after subtracting the number of volunteers from the annual military quota, in order to meet the remainder of the conscription quota (target) assigned to that particular colonial district. We have shown that the colonial-district-level share (lottery rate) variable observed in our individual archive dataset tracks closely the colonial-district-level share from a secondary archive source from the National Archives of Senegal in Dakar.

[58]The second sub-step ensures that the simulated colony-district-level share of reservists would be the same as the actual colony-district-level share of reservists, while only keeping the assignment into the regular soldier camp or the military reservist camp random at the individual level.

[59]Namely we only keep the localities that have a maximum matching score between 0.8 and 1 as returned by the STATA *matchit* command, which represents around 77% of the total individual sample. We will perform more sensitivity check by perhaps including localities with lower matching scores (between 0.5 and 0.8 for instance) in the future.

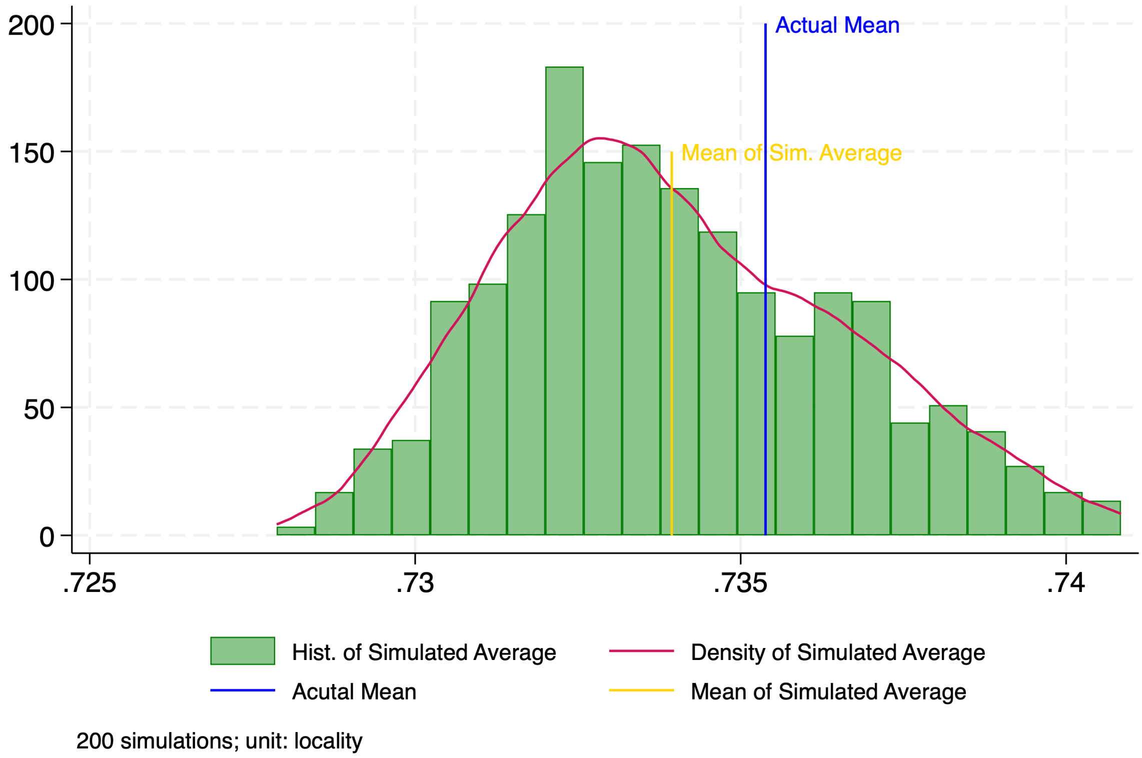
seems to be a higher peak around the sample mean for the simulated distributions compared to the observe one, the distributional spreads are quite similar between simulation results and the observed share density. This observation is reassuring as it suggests that the lottery system was not subject to mean reversion, indicating that the observed distribution likely resulted from a quasi-random experiment.

Figure C2: Monte Carlo simulation results: histogram (orange) of the share of reservists in the archive dataset across localities, and kernel densities of the simulated distribution of the share of reservists. The kernel density of the actual distribution of the share of reservists is colored in blue. Unit: locality.



While the similarity in spread between the actual and simulated distributions of reservist shares is reassuring, the greater concentration around the mean in the simulated data may raise concerns. This is further examined in Figure C3, which displays the actual mean of the observed share distribution (the blue bar), versus the averages of simulated shares across 200 simulations (represented by the green histogram). In this graph, we observe that although the actual mean (blue line) slightly differs from the mean of the simulated averages (orange line), the observed mean still falls within the middle range of the simulated average distribution. And across the 200 simulations conducted, we could only reject the null that the simulated mean and the observed mean are the same at the 5% significance level, 4% (8/200) of the time.

Figure C3: Monte Carlo simulation results: histogram (green) of the simulated averages over 200 simulations. In blue the observed mean of reservist in the data and in orange the mean of simulated averages. Unit: locality.



A.9 Robustness checks

Table C5: Relationship between schooling infrastructure and share of reservists by colonial-district-level military forced labor activation rate - robustness check

	Community Schools							
	Nb schools		Student to child ratio		Teacher to child ratio		Student to teacher ratio	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.00	0.05	0.08	-0.70	0.00	-0.02	0.26	-2.91
	(0.07)	(0.21)	(0.24)	(0.67)	(0.01)	(0.02)	(2.04)	(11.66)
Share_Reservists x District Activation Rate		0.12		-1.57		-0.06		-6.29
		(0.41)		(1.37)		(0.04)		(22.19)
Mean DV	0.33	0.33	0.44	0.44	0.02	0.02	34.37	34.37
St. Dev. DV	0.75	0.75	3.79	3.79	0.16	0.16	15.35	15.35
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	3710	3710	3710	3710	1519	1519
Adj. R^2	0.07	0.07	0.01	0.01	0.01	0.01	0.14	0.14
	Public Schools							
	Nb schools		Student to child ratio		Teacher to child ratio		Student to teacher ratio	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.11	0.62	-1.26	0.84	-0.02	0.01	1.60	-8.22
	(0.21)	(1.09)	(1.24)	(1.71)	(0.02)	(0.03)	(2.31)	(8.88)
Share_Reservists x District Activation Rate		1.54		4.23		0.06		-19.98
		(1.99)		(5.55)		(0.10)		(17.32)
Mean DV	0.77	0.77	1.25	1.25	0.03	0.03	46.30	46.30
St. Dev. DV	2.36	2.36	12.79	12.79	0.23	0.23	19.33	19.33
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5607	5607	3710	3710	3710	3710	2436	2436
Adj. R^2	0.02	0.02	0.00	0.00	0.00	0.00	0.08	0.08

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. District fixed effects are colonial fixed effects. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table C6: Main results excluding North districts

Census 1976

	Fertility		Primary school		Eligible Men Pop			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.00	0.09*	0.01	-0.08***	3.53	-13.36		
	(0.01)	(0.05)	(0.01)	(0.03)	(7.75)	(30.13)		
Share_Reservists x District Activation Rate		0.24*		-0.23***		-45.08		
		(0.13)		(0.08)		(82.43)		
Mean DV	0.44	0.44	0.08	0.08	73.47	73.47		
St. Dev. DV	0.14	0.14	0.08	0.08	132.15	132.15		
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes		
Obs	4146	4146	4151	4151	5468	5468		
Adj. R^2	0.12	0.12	0.07	0.08	0.12	0.12		

Census 1987

	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.12*	0.00	-0.06**	-0.02	0.03	-0.00	-0.18
	(0.02)	(0.07)	(0.01)	(0.02)	(0.03)	(0.13)	(0.03)	(0.12)
Share_Reservists x District Activation Rate		0.29*		-0.16***		0.12		-0.46
		(0.17)		(0.06)		(0.31)		(0.31)
Mean DV	0.89	0.89	0.06	0.06	-0.14	-0.14	0.33	0.33
St. Dev. DV	0.22	0.22	0.08	0.08	0.41	0.41	0.38	0.38
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5468	5468	5468	5468	5468	5468	5468	5468
Adj. R^2	0.17	0.17	0.08	0.08	0.21	0.21	0.06	0.06

Census 1998

	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.00	0.04	0.00	-0.09***	0.00	-0.02	0.02	-0.05
	(0.01)	(0.06)	(0.01)	(0.03)	(0.02)	(0.09)	(0.02)	(0.10)
Share_Reservists x District Activation Rate		0.09		-0.25***		-0.07		-0.18
		(0.14)		(0.08)		(0.24)		(0.28)
Mean DV	0.82	0.82	0.09	0.09	-0.30	-0.30	0.17	0.17
St. Dev. DV	0.18	0.18	0.11	0.11	0.26	0.26	0.31	0.31
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5468	5468	5468	5468	5468	5468	5468	5468
Adj. R^2	0.10	0.10	0.14	0.14	0.11	0.11	0.12	0.12

Census 2009

	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.02	0.02	-0.02	-0.12**	-0.01	0.02	-0.00	0.02
	(0.02)	(0.08)	(0.01)	(0.05)	(0.02)	(0.10)	(0.02)	(0.08)
Share_Reservists x District Activation Rate		0.02		-0.28**		0.08		0.05
		(0.18)		(0.13)		(0.27)		(0.23)
Mean DV	0.94	0.94	0.23	0.23	-0.41	-0.41	0.29	0.29
St. Dev. DV	0.27	0.27	0.16	0.16	0.32	0.32	0.29	0.29
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5466	5466	5468	5468	5468	5468	5468	5468
Adj. R^2	0.02	0.02	0.23	0.24	0.23	0.23	0.21	0.21

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. For census 1976: Schools, Dispensaries, Markets indicate the presence of the infrastructure in the locality. For census 1987 and 2009: Fertility is measured as the ratio of children between 0 and 5 over the number of women 15-50 in the village. Primary school is measured as the share of the population in the locality aged 6 years or more, having attained at least primary school; Wealth Index is the average of the wealth index computed using principal component analysis of the housing characteristics of the households within the locality; Interquartile range is the difference between the wealth index of the 10 percentile and the 90th percentile in the village. District fixed effects are colonial fixed effects. Discrepancies in the number of observations for Gini is due to insufficient number of households in the locality to calculate Gini. North districts include Bourem, Gao, Goundam, Gourma, Tombouctou. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table C7: Main results accountimg for potential spatial spillovers

	Census 1976							
	Fertility		Primary school		Eligible Men Pop			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	-0.00	0.07	0.00	-0.09***	0.71	-35.34		
	(0.01)	(0.05)	(0.01)	(0.03)	(9.85)	(50.64)		
Share_Reservists x District Activation Rate		0.21*		-0.25***		-101.85		
		(0.12)		(0.08)		(149.39)		
Neighbors : Share_Soldiers	0.01	-0.03	0.01	-0.03	0.04	17.75		
	(0.02)	(0.04)	(0.01)	(0.02)	(13.73)	(42.38)		
Neighbors : Share_Reservists x District Activation Rate		-0.11		-0.10		50.59		
		(0.10)		(0.06)		(117.10)		
Mean DV	0.44	0.44	0.08	0.08	99.65	99.65		
St. Dev. DV	0.14	0.14	0.08	0.08	147.24	147.24		
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes		
Obs	4248	4248	4253	4253	4268	4268		
Adj. R^2	0.11	0.12	0.08	0.08	0.08	0.08		

	Census 1987							
	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.13*	0.00	-0.07***	-0.03	-0.05	-0.01	-0.18
	(0.02)	(0.07)	(0.01)	(0.02)	(0.03)	(0.12)	(0.03)	(0.12)
Share_Reservists x District Activation Rate		0.32*		-0.19***		-0.06		-0.47
		(0.17)		(0.06)		(0.31)		(0.31)
Neighbors : Share_Soldiers	0.02	0.06	0.00	0.02	0.01	0.21**	-0.03	-0.02
	(0.02)	(0.05)	(0.01)	(0.02)	(0.03)	(0.10)	(0.03)	(0.10)
Neighbors : Share_Reservists x District Activation Rate		0.09		0.04		0.52**		0.04
		(0.12)		(0.04)		(0.24)		(0.23)
Mean DV	0.89	0.89	0.06	0.06	-0.15	-0.15	0.34	0.34
St. Dev. DV	0.22	0.22	0.08	0.08	0.43	0.43	0.39	0.39
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5573	5573	5573	5573	5573	5573	5573	5573
Adj. R^2	0.17	0.17	0.08	0.08	0.26	0.26	0.09	0.09

	Census 1998							
	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.05	-0.00	-0.10***	0.00	-0.01	0.02	-0.06
	(0.01)	(0.05)	(0.01)	(0.03)	(0.02)	(0.08)	(0.02)	(0.10)
Share_Reservists x District Activation Rate		0.13		-0.27***		-0.04		-0.21
		(0.14)		(0.08)		(0.23)		(0.27)
Neighbors : Share_Soldiers	-0.01	-0.01	0.00	0.02	-0.02	0.03	-0.00	0.05
	(0.02)	(0.05)	(0.01)	(0.03)	(0.02)	(0.06)	(0.03)	(0.07)
Neighbors : Share_Reservists x District Activation Rate		0.01		0.06		0.13		0.15
		(0.12)		(0.06)		(0.15)		(0.21)
Mean DV	0.82	0.82	0.09	0.09	-0.31	-0.31	0.17	0.17
St. Dev. DV	0.18	0.18	0.10	0.10	0.25	0.25	0.31	0.31
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5573	5573	5573	5573	5573	5573	5573	5573
Adj. R^2	0.10	0.10	0.14	0.14	0.13	0.13	0.12	0.12

	Census 2009							
	Fertility		Primary school		Wealth Index		Interquartile range	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share_Soldiers	0.01	0.02	-0.02*	-0.14***	-0.01	-0.01	-0.00	0.01
	(0.02)	(0.07)	(0.01)	(0.05)	(0.02)	(0.09)	(0.02)	(0.08)
Share_Reservists x District Activation Rate		0.01		-0.31**		0.02		0.03
		(0.17)		(0.12)		(0.25)		(0.22)
Neighbors : Share_Soldiers	-0.05***	-0.11**	-0.02	-0.01	-0.02	0.04	0.01	0.10
	(0.02)	(0.06)	(0.01)	(0.04)	(0.02)	(0.07)	(0.02)	(0.07)
Neighbors : Share_Reservists x District Activation Rate		-0.15		0.03		0.16		0.23
		(0.13)		(0.09)		(0.17)		(0.18)
Mean DV	0.94	0.94	0.23	0.23	-0.41	-0.41	0.29	0.29
St. Dev. DV	0.27	0.27	0.16	0.16	0.31	0.31	0.29	0.29
District-group FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	5571	5571	5573	5573	5573	5573	5573	5573
Adj. R^2	0.02	0.02	0.23	0.24	0.24	0.24	0.21	0.21

Notes: Regressions at the locality level with sample restricted to the panel of matched localities across censuses. For census 1976: Schools, Dispensaries, Markets indicate the presence of the infrastructure in the locality. For census 1987 and 2009: Fertility is measured as the ratio of children between 0 and 5 over the number of women 15-50 in the village. Primary school is measured as the share of the population in the locality aged 6 years or more, having attained at least primary school; Wealth Index is the average of the wealth index computed using principal component analysis of the housing characteristics of the households within the locality; Interquartile range is the difference between the wealth index of the 10 percentile and the 90th percentile in the village. District fixed effects are colonial fixed effects. Discrepancies in the number of observations for Gini is due to insufficient number of households in the locality to calculate Gini. Neighbors include all localities within a radius of 5km. Standard errors clustered at the locality levels for 1976 and locality-year level for the 1987 and 2009 regression. Significance levels: *p<0.1; **p<0.05; ***p<0.01