

Online Appendix A: Geography of referrals

We know the place of attestation (where the recruit officially enlisted). We do not have information where soldiers were stationed subsequently.

The practice of referrals was not very clustered by military base. It was equally common in Kumasi and Accra, the two main and largest military bases (Table A1). Referrals were more often used in the Northern Territories (NT) though not unanimously; there are exceptions such as Wa and Gambaga. Geographic locations of the military bases are shown in Figure A1.

In the paper we use ethnic group fixed effects as a control in the regressions of observables (height and chest circumference) and unobservables (desertions and discharged as medically unfit/inefficient). Ethnicity is a very disaggregate geographic, economic and community marker. Alternatively, one could use the 31 administrative districts of Ghana and the 13 countries of origin for areas outside of Ghana, but this is a HIGHER aggregate. Moreover, districts are drawn for organisational purposes by the coloniser. As a result, districts likely comprehend *more* heterogeneous communities than ethnic groups. Moreover, we were not able to identify the longitude/latitude of every place of birth. Using ethnic groups avoids attrition.

Table A1: Referrals by place of attestation

Place of attestation	Region	N recruits	Percentage referrals
Kumasi	Ashanti	3,932	32.3%
Accra	Gold Coast	1,630	30.3%
Tamale	NT	749	66.9%
Wa	NT	361	1.9%
Lome	Togo	354	56.8%
Cape Coast	Gold Coast	231	10.8%
Zuaragu	NT	197	62.4%
Gambaga	NT	181	23.2%
Mpraeso	Gold Coast	159	3.1%
Sekondi	Gold Coast	146	0.7%
Missing information		124	45.2%
Other		396	42.7%
Total		8640	34.2%

Note: "Other" includes 33 places. None of those places has more than 51 recruits.

Figure A1: Geographic location of attestation places

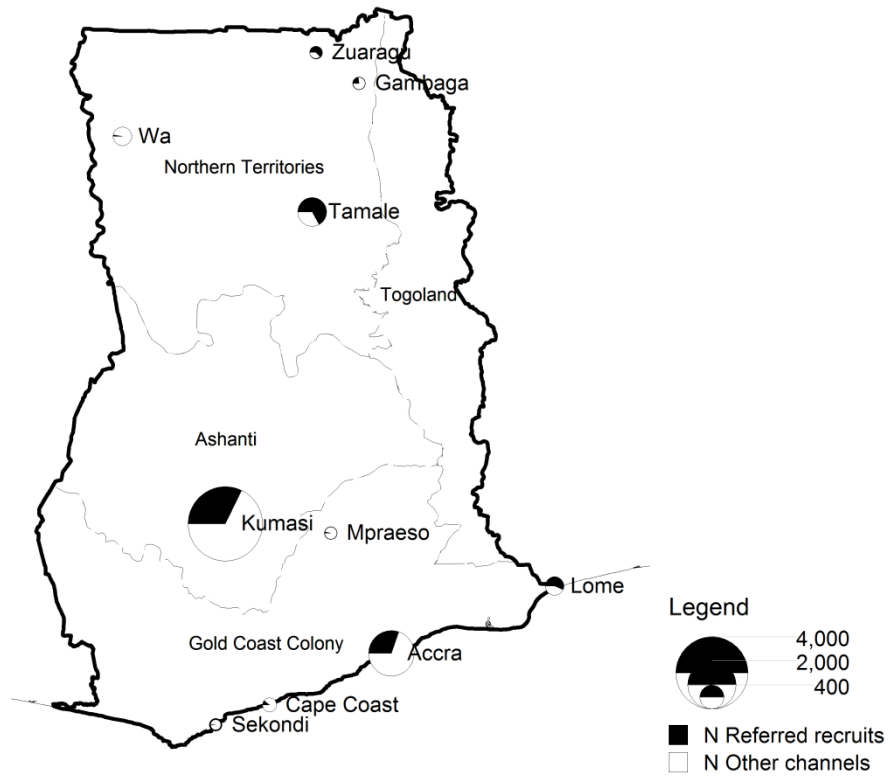
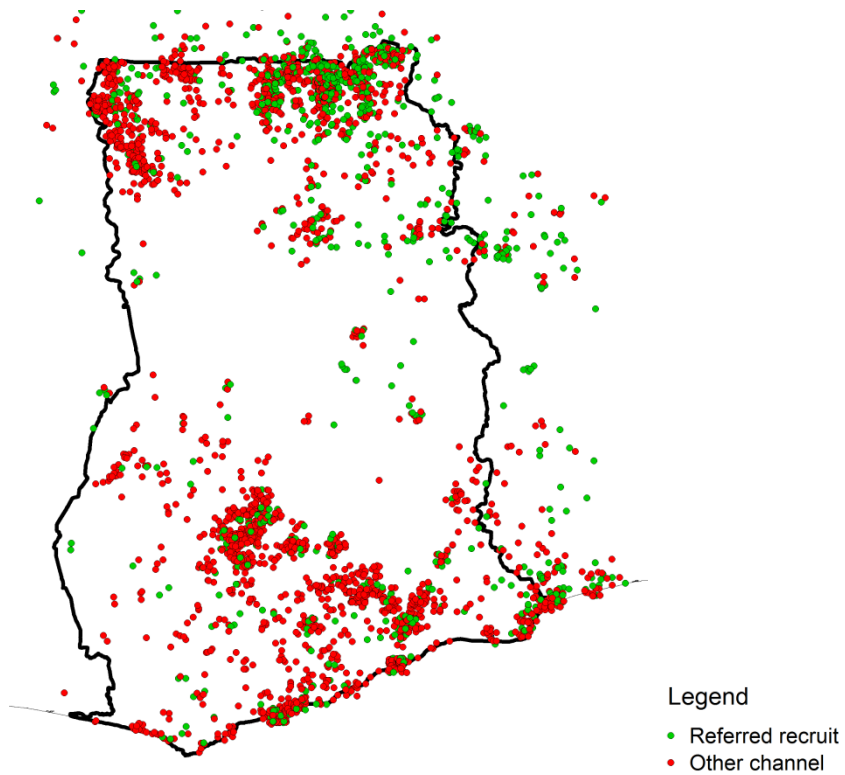


Figure A2 Referrals by place of birth



Note: Dots represent recruit's place of birth and were jittered.

Online Appendix B: Order of referrals

There may have been more pressure on higher ranking soldiers to generate more recruits. If this is the case, the soldier would have to go down lower and lower on the quality distribution within his network to generate more recruits. This would mean on average their recruits were of lower quality, even if they were attempting to screen for the army and not acting opportunistically. Our main robustness check was eliminating soldiers who make 9 or more referrals, but the deterioration of quality with quantity may occur at a much earlier stage.

This suggests that the order of referral is important (with the first referred recruit being better because better known to the referee than the second referral and so on). From our data set, we can extract information on the order of referrals, because we know the time of enlistment of the referred recruits as well as by whom the recruit was referred.

In calculating the order of the referral there is one complication however: we have a truncated sample in that we do not observe any enlistments of bringers AND referrals before 1908. Hence, we need to limit the variable to bringers who enlisted after 1908 (so that all of his subsequent referrals are observed in our sample).

We then test the order of the referral as an additional variable in the regressions. The order of the referral is indeed significant. The observable regression turns out as hypothesised. Height and chest circumference of later referrals decreased. In terms of unobservable characteristics, the order of referral has a negative but insignificant correlation with desertion and positive impact on dismissals. See regression output below. Interaction terms with the rank of the referring soldiers are never significant (not shown). However, it does not change any of our main results.

Table 5c: Referral and observable quality -- with controls & order of referral

	Height (cm)		Chest Circumference (cm)	
	(1)	(2)	(1)	(2)
Recruiting channels				
Referred by fellow soldier	0.909*** (0.329)	0.451* (0.237)	0.304* (0.178)	-0.136 (0.136)
Order of referral (1,2,...)	-0.100** (0.048)	-0.061 (0.046)	-0.196*** (0.033)	-0.156*** (0.034)
Sent by traditional chief	0.396 (0.468)	0.097 (0.267)	0.666* (0.346)	0.416 (0.363)
Volunteer	1.132** (0.457)	1.093*** (0.354)	0.914** (0.358)	0.819** (0.311)
Relatives in the army				
Recruit has relative in the army (1=yes)	0.306 (0.507)	-0.692** (0.266)	0.917*** (0.301)	0.205 (0.185)
Bringer is kin (1=yes)	-0.921 (0.829)	-0.347 (0.763)	-0.754 (0.579)	-0.394 (0.542)
Bringer(1=yes)	0.254 (0.697)	0.109 (0.656)	0.522 (0.370)	0.498 (0.367)
Previous occupation				
Farmer	0.149 (0.516)	-0.682 (0.435)	0.072 (0.377)	-0.465 (0.349)
Armed forces (police, army)	1.038 (0.649)	0.475 (0.498)	1.238** (0.479)	0.758* (0.434)
Literate	0.393 (0.353)	0.295 (0.330)	-0.458** (0.210)	-0.517** (0.201)
Skill index FE (1 unskilled, 5 highly skilled)	YES	YES	YES	YES
Other controls				
Motor Transport Unit (1=yes)	0.725** (0.360)	1.606*** (0.283)	0.424 (0.268)	0.532* (0.270)
Year of enlistment FE	YES	YES	YES	YES
Age fixed effects	YES	YES	YES	YES
Ethnic group FE		YES		YES
Observations	7,827	7,827	7,682	7,682
R-squared	0.093	0.159	0.148	0.197
N_clust	81	81	81	81

Estimator is OLS. Standard errors in (1) clustered by ethnic groups.

Table 6c: Referral and observable quality -- with controls & order of referral

	Desertion		Dismissal	
	(1)	(2)	(1)	(2)
Recruiting channels				
Referred by fellow soldier	1.291** (2.324)		0.954 (-0.519)	
Rank of bringer at the time of referral				
Low rank (Private, Lance Corporal, Corporal)		1.116 (0.839)		0.962 (-0.372)
High rank (Sergeant, CSM, RSM)		1.419*** (3.011)		0.947 (-0.537)
Unknown				
Order of referral (1,2, ...)	0.970 (-1.111)	0.971 (-1.134)	1.027* (1.865)	1.027* (1.864)
Sent by traditional chief	1.490*** (3.271)	1.497*** (3.309)	1.385*** (3.058)	1.385*** (3.051)
Volunteer	1.040 (0.224)	1.049 (0.270)	1.088 (0.529)	1.088 (0.526)
Bringer (yes=1)	1.120 (0.418)	1.124 (0.429)	0.890 (-0.587)	0.890 (-0.590)
Relatives in the army				
Recruit has relative in the army (yes=1)	1.045 (0.385)	1.030 (0.258)	1.185** (1.995)	1.186** (2.002)
Bringer is kin (yes=1)	0.975 (-0.134)	1.046 (0.235)	1.345** (2.014)	1.340** (1.970)
Previous occupation:				
Literate (yes=1)	0.802 (-1.274)	0.798 (-1.305)	0.964 (-0.256)	0.964 (-0.252)
Farmer	1.115 (0.245)	1.112 (0.239)	2.235 (1.442)	2.236 (1.443)
Armed forces (police, army)	0.917 (-0.265)	0.899 (-0.322)	1.085 (0.343)	1.086 (0.345)
Skills in previous occupation fixed effects	YES	YES	YES	YES
Other controls				
Motor Transport Unit (yes=1)	0.679* (-1.764)	0.689* (-1.700)	0.823 (-0.995)	0.822 (-0.999)
Height (cm)	1.006 (1.206)	1.006 (1.198)	1.004 (0.765)	1.004 (0.767)
Chest circumference (cm)	1.009 (1.255)	1.009 (1.232)	0.985** (-2.252)	0.985** (-2.246)
Year of enlistment fixed effects	YES	YES	YES	YES
Year of service fixed effects	YES	YES	YES	YES
Age at enlistment fixed effects (14-23 yrs)	YES	YES	YES	YES
Ethnic group fixed effects	YES	YES	YES	YES
Chi2-Test of equality in hazards				
H0: High rank=low rank (p-value)		0.03		0.92
Number of failures (desertions/dismissals)	843	843	1187	1187
Number of subjects	7134	7134	7134	7134

Cox proportional hazard model. Reported coefficients are hazard ratios; z-values in parentheses; standard errors clustered by ethnic groups.

*** p<0.01 ** p<0.05 * p<0.1

Online Appendix C: Unknown date of discharge

The survival analysis requires information of time at risk. Unfortunately, a substantial proportion of attestation papers (38%) are silent on termination. After the war, the army modified its system for recording details of discharge. While desertion and dismissals were still recorded on attestation papers, soldiers who completed their service normally received distinct discharge papers.

In our analysis we assumed a missing date of discharge to imply that the soldier served at least until 15th November 1918. The 15th November is the earliest known discharge date within the GCR after the war ended on 11th November 1918. As demobilisation lasted until early 1919 (Killingray 1982: 380), some soldiers may have been a longer time at risk than assumed.

This is a reasonable assumption bringing the number of soldiers serving during WWI in line with various estimates reported by Thomas (1975) and Killingray (1982). We nevertheless worry that this assumption may be incorrect. To investigate whether this affects our results, we reestimate the desertion and dismissal models without those recruits for whom the circumstances of contract termination are not recorded. This, however, raises another potential problem, namely, selection bias induced by missing data. The switch in recording dealt with WWI and demobilisation. A cursory inspection of the data indicates a strong time trend in the proportion of missing observations, with less information available for individuals recruited towards the end of our study period. Moreover, recruits whose terms ended as soon as the war ended were more likely to receive distinct discharge papers. This suggests using the term of enlistment “Duration of war” as instrument to correct for possible selection bias.

To this effect, we reestimate the duration models with a selection correction using the method proposed by Boehmke, Morey and Shannon (2006). This model assumes a Weibull distribution for the hazard. Our data easily accommodates a Weibull distribution, so this is unproblematic. Results are reported in Table C below. Results with the selection correction confirm our findings, with recruits referred by high ranking fellow soldiers more likely to desert. Other results are confirmed as well.

Table C: Referral and unobservable quality -- with selection correction

	Desertion				Dismissal			
	(1)		(2)		(1)		(2)	
	Coef.	z	Coef.	z	Coef.	z	Coef.	z
Duration equation								
<i>Referred by fellow soldier</i>	0.260***	2.81			0.005	0.04		
Rank of bringer:								
Low rank (Pte, L Cpl, Cpl)			0.094	0.67			0.060	0.52
High rank (Sgt, CSM, RSM, unknown)			0.383***	3.41			0.058	0.58
Other recruiting channels								
Sent by traditional chief	0.401***	2.81	0.348***	2.36	0.398***	3.70	0.396	4.52
Volunteer	0.164	0.90	0.157	0.70	-0.015	-0.08	0.041	0.26
Relatives in the army								
Recruit has relative in the army			-0.234	-1.47			0.212**	2.37
Bringer is kin (1=yes)			0.218	0.97			0.196*	1.78
Bringer (1=yes)			-0.856**	-3.09			-0.576**	-2.21
Previous occupation:								
Farmer			0.559***	2.84			0.445***	3.52
Armed forces (police, army)			-0.136	-0.44			0.126	0.37
Literate			-0.231	-1.10			-0.051	-0.32
Skills in previous occupation FE			YES				YES	
Other controls								
Motor Transport Unit			0.341	1.50			.546***	2.42
Height (cm)			0.001	0.33			0.004	0.97
Chest circumference (cm)			0.008	0.88			-0.016**	-2.34
Age of recruit at enlistment FE	YES		YES		YES		YES	
Year of enlistment FE	YES		YES		YES		YES	
Intercept	-7.444	-11.41	-7.587***	-6.73	-7.436***	-32.51	-8.536***	-9.36
Selection equation								
Referred by fellow soldier	0.075	0.85	0.140	1.55	0.140	1.56	0.240***	2.64
Sent by traditional chief	0.467***	2.92	0.520***	3.20	0.521***	3.20	0.628***	3.04
Volunteer	-0.035	-0.34	-0.016	-0.16	-0.017	-0.17	0.116	1.17
Term: Duration of war	0.300**	2.070	0.316**	2.21	0.313**	2.24		
Motor Transport Unit	-1.148***	-8.74	-1.113***	-8.25	-1.112***	-8.25	-0.730***	-5.57
Year of enlistment FE	YES		YES		YES		YES	
Intercept	1.270***	4.37	-0.172	-1.08	-0.1709	-1.08	130.84***	5.31
Chi2-Test of equality in hazards (p-value)			0.047				0.984	
Number of uncensored observations	5271		4786		5271		4786	
Number of observations	8460		7975		8460		7975	

Estimator is Weibull duration model with selection correction; standard errors clustered by ethnic groups.

*** p<0.01 ** p<0.05 * p<0.1