Refugees and Deforestation in sub-Saharan Africa

Supplemental Appendix
April 24th, 2021

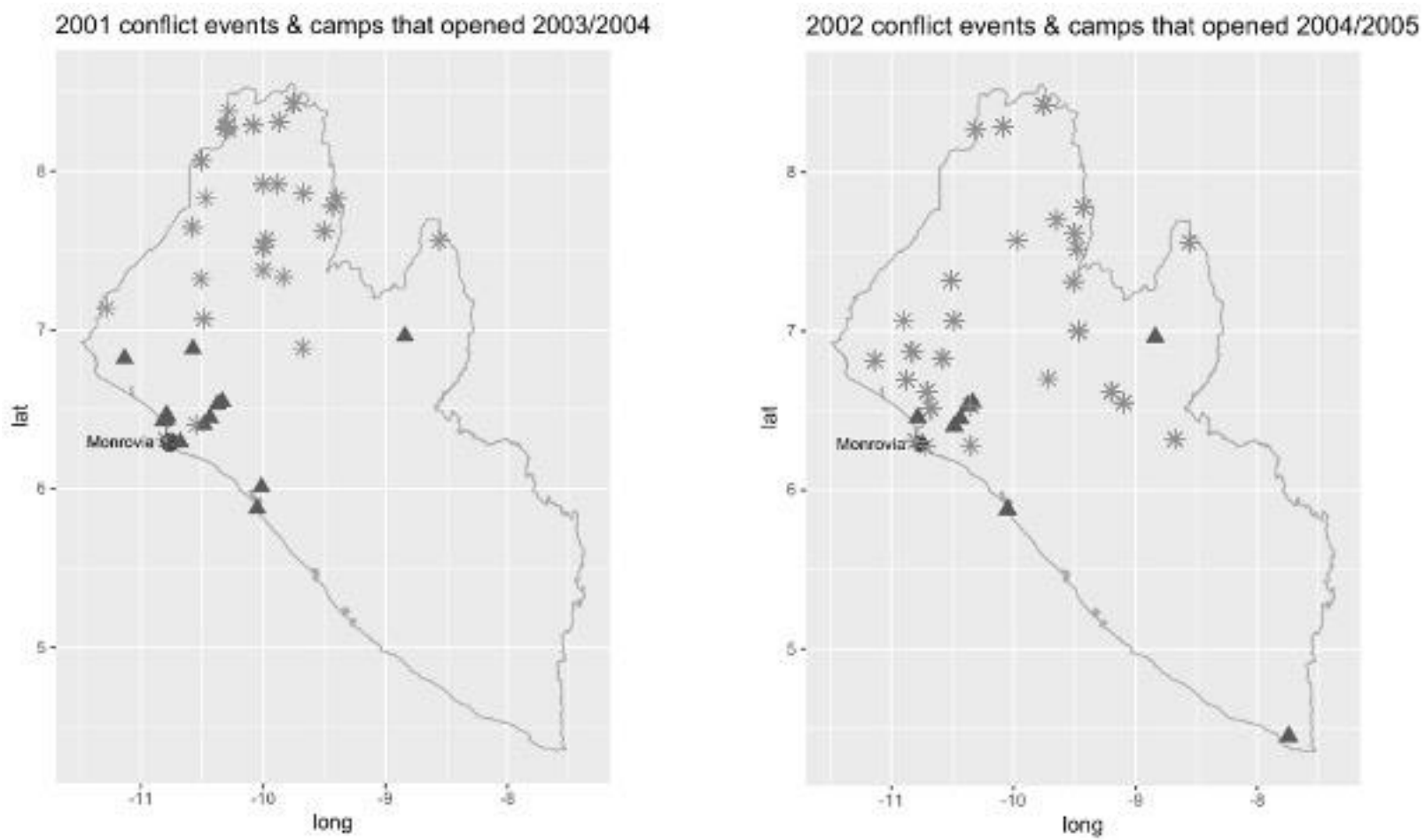
Appendix A1: Examining noise in forest cover data for camp areas in Liberia

While examining the data, I noticed that areas near camps in Liberia often transitioned to zero forest cover (according to the GFC data) a few years prior to camp creation. For this reason, I omit Liberian camps from the study. In this section, I explain why I think the camp data for Liberia may be noisier than other countries.

My initial assumption was that civil conflict during the war resulted in delays in humanitarian programming. UNHCR reported in 2001 and 2003 that it was unable to reach refugees living in war-affected areas of the country (Kamara 2003; UNHCR 2001, 2003). But upon closer inspection, I find that the areas around camps that exhibit early transitions to zero forest cover are primarily clustered around the capital. And while some of these camps went under attack during the war, this geographically does not correspond with the areas where most of the fighting took place in 2001 and 2002: the north and northwestern areas of the country (see Figure A4).

A more likely possibility is that the locations registered as refugee camps following the ARD protocol either served as settlements for internally displaced persons prior to the refugee camp designation (International Rescue Committee 2003; news24 2003), or were established close to pre-existing IDP camps (Marie 2003). The congregating of IDPs near the capital city would be consistent with Alix-Garcia, Bartlett, and Saah (2013), who find that internally displaced people in Darfur congregated outside of cities for protection and contributed significantly to declines in vegetation in these areas.

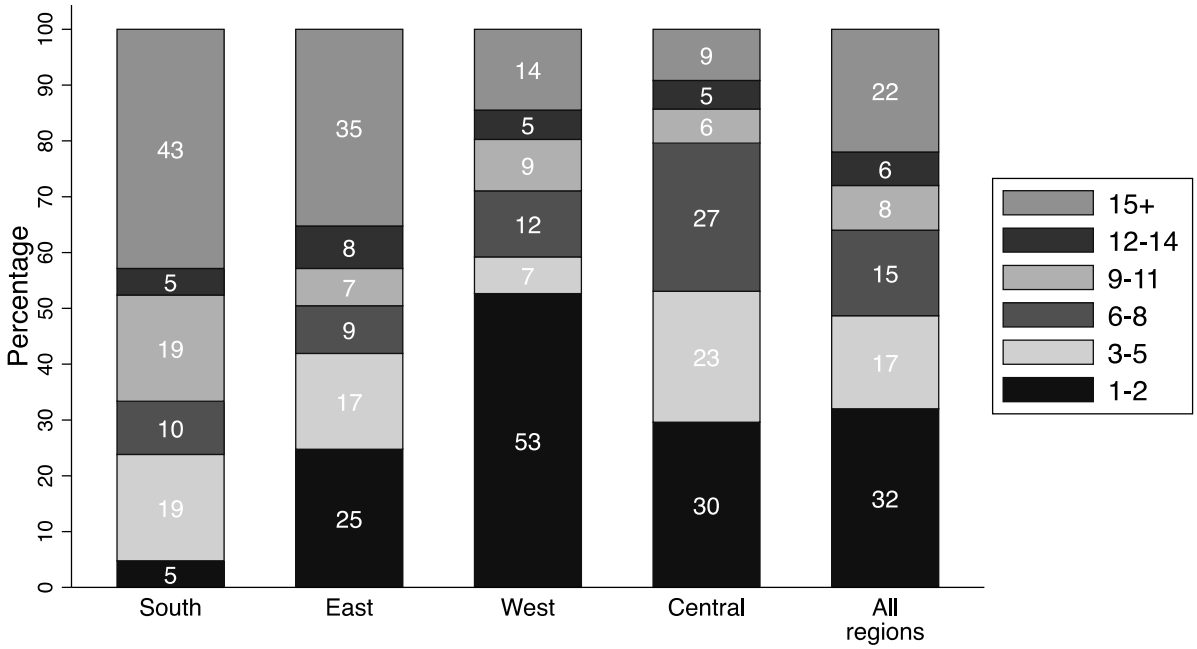
Figure A1.1: map of conflict events and camp openings in Liberia 2-3 years after



Source: authors' production using the Armed Conflict Location and Event Data (ACLED) (Raleigh et al. 2010) and the ARD data for Liberia. I restrict my ACLED data points to the following event categories: battles, explosions/remote violence, and violence against civilians. Conflict events are marked with a red asterisk, camps are marked with a blue triangle, and the capital, Monrovia, is a dark green circle.

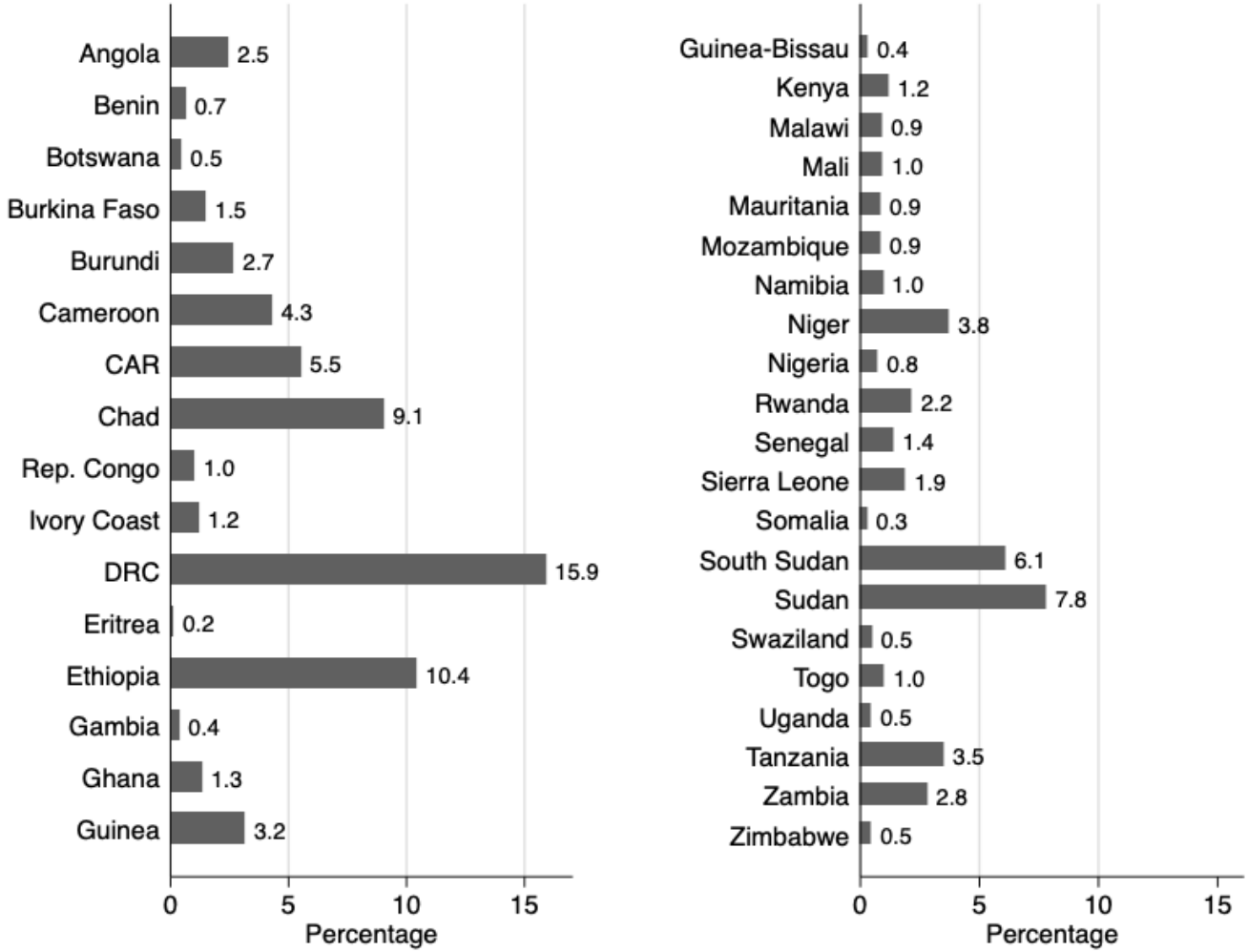
Appendix A2: Additional descriptive figures

Figure A2.1: Number of years camps operated by region



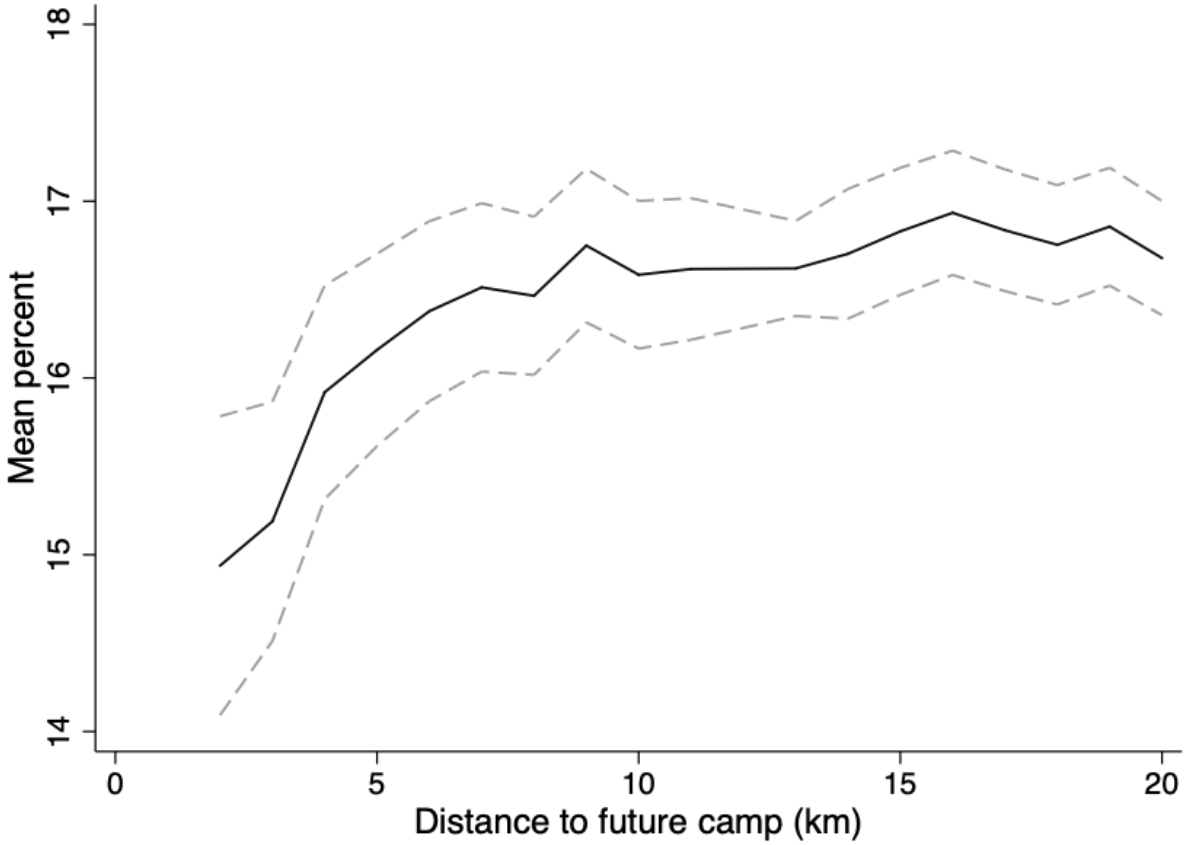
Source: author's calculations using ARD data. Regional categorizations of countries are based on the African Union's official classification.

Figure A2.2: distribution of tiles by country (percentage)



Source: author's calculations based on 0.01° sample tiles and FAO's ADM0 shapefile.

Figure A2.3: Percentage forest cover in 2000 by distance to a future camp (camp is not open as of 2000)

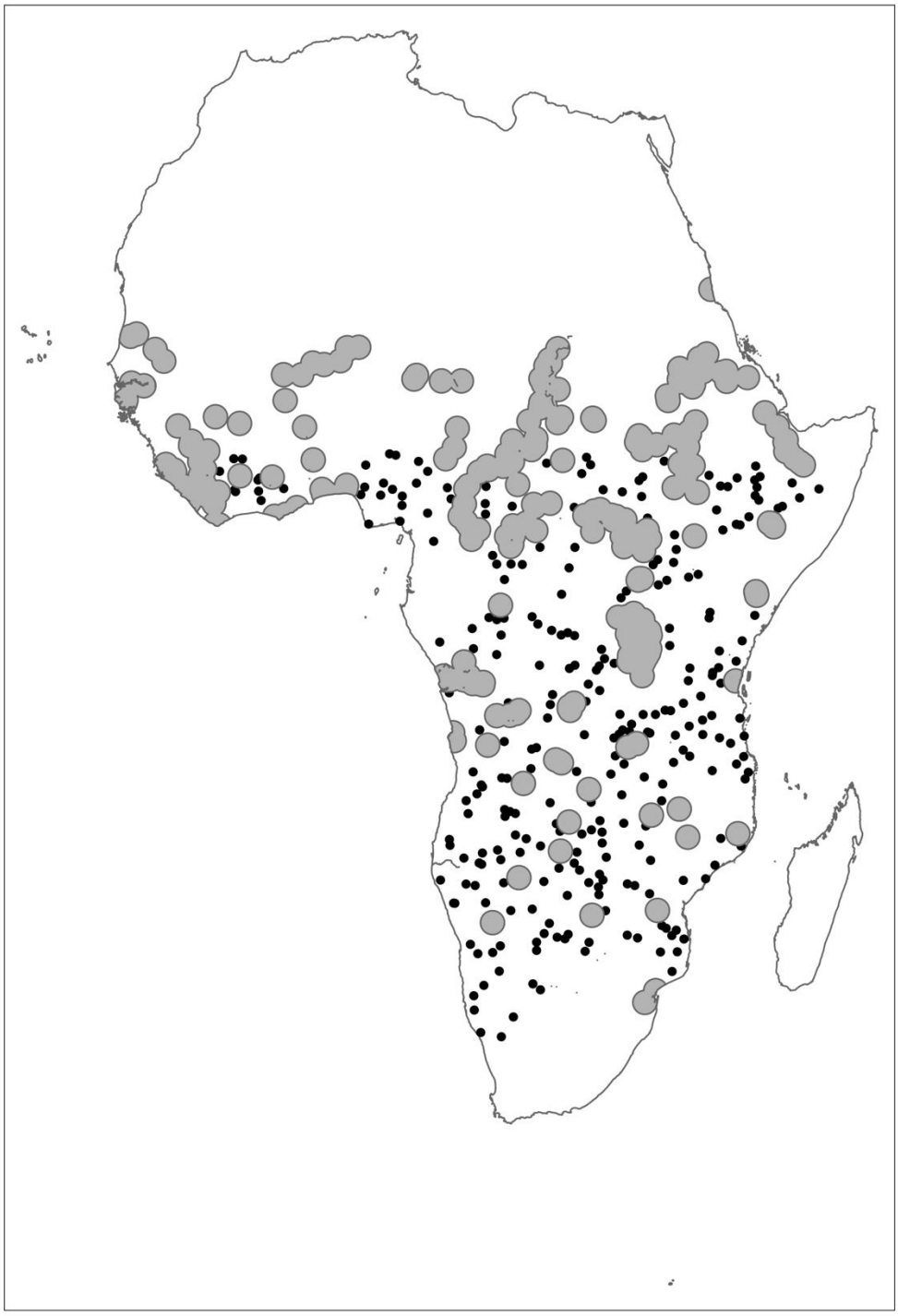


Source: author's calculations using ARD and GFC data. Plot constructed using sample of 0.01° tiles. Dotted lines indicated 95% confidence intervals around the mean. Sample restricted to pixels within 20 km of a future camp not exposed to a refugee camp in 2000.

Appendix A3: Method to sample comparison group from outside of camp areas

As a robustness check, and out of concern of possible contamination of the comparison group, I collect zonal statistics for a new sample of tiles selected exclusively from areas far from the camps. To do this, I generated a map of Africa south of the Sahara Desert and excluding islands and South Africa. From this map, I removed all areas within 100 km of an ARD refugee camp active during the study period. I then randomly sampled 273 points from the remaining areas. I created a 30 km buffer around each of these points and melted overlapping buffers together. I defined 0.01° tiles within these buffers and calculated zonal statistics for each tile. Figure A3.1 maps out the locations of the points used to produce this comparison sample and highlights the 100 km areas around camps that were omitted from selection.

Figure A3.1: location of points used to produce comparison sample

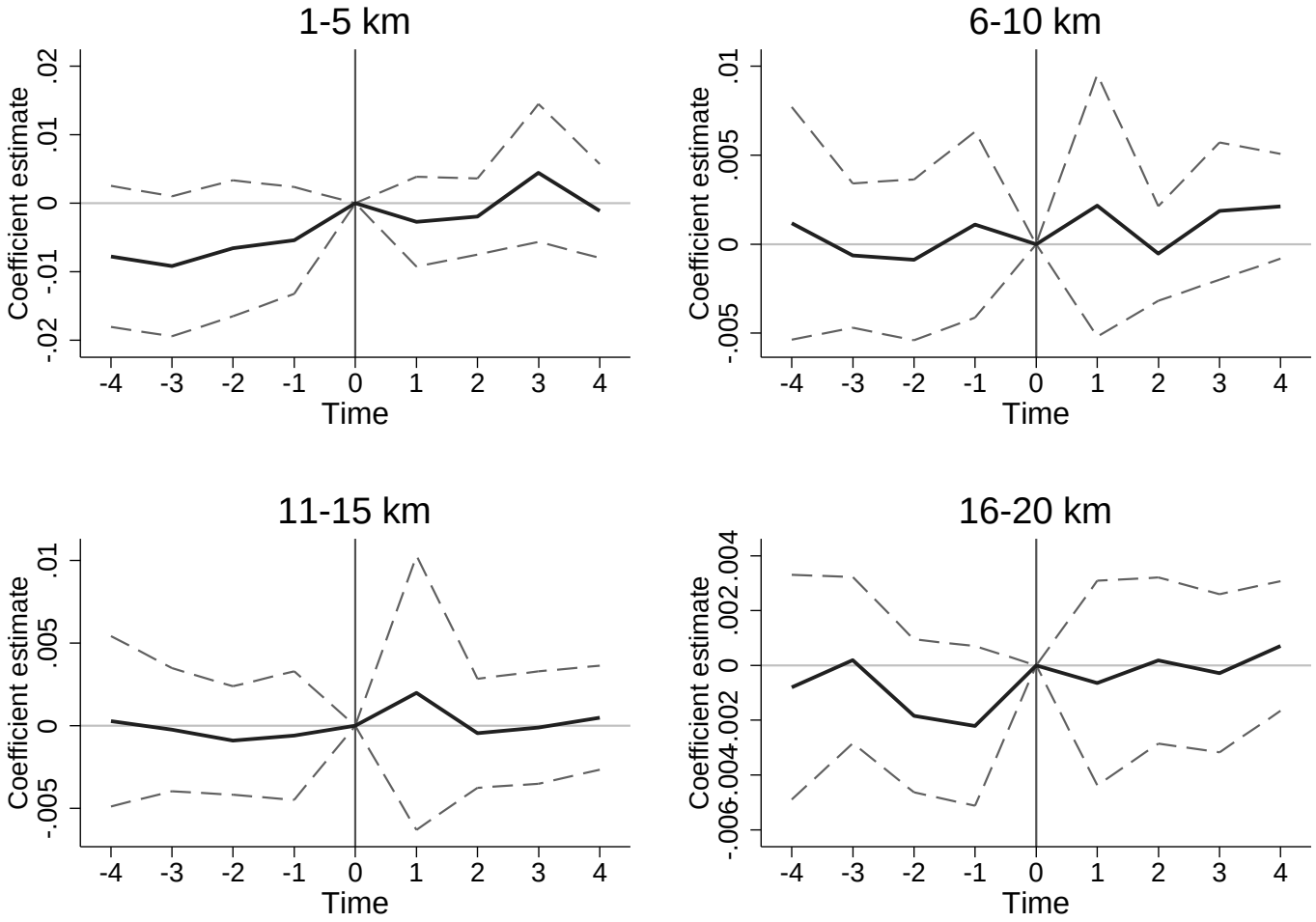


Source: author's calculations using ARD data. Grey regions represent the 100 km area around ARD camps used in the study: these are omitted when selecting points. The points used to build the comparison sample are shown in black. Map uses Eckert 4 projection.

Appendix A4: Event study plots not shown in main text

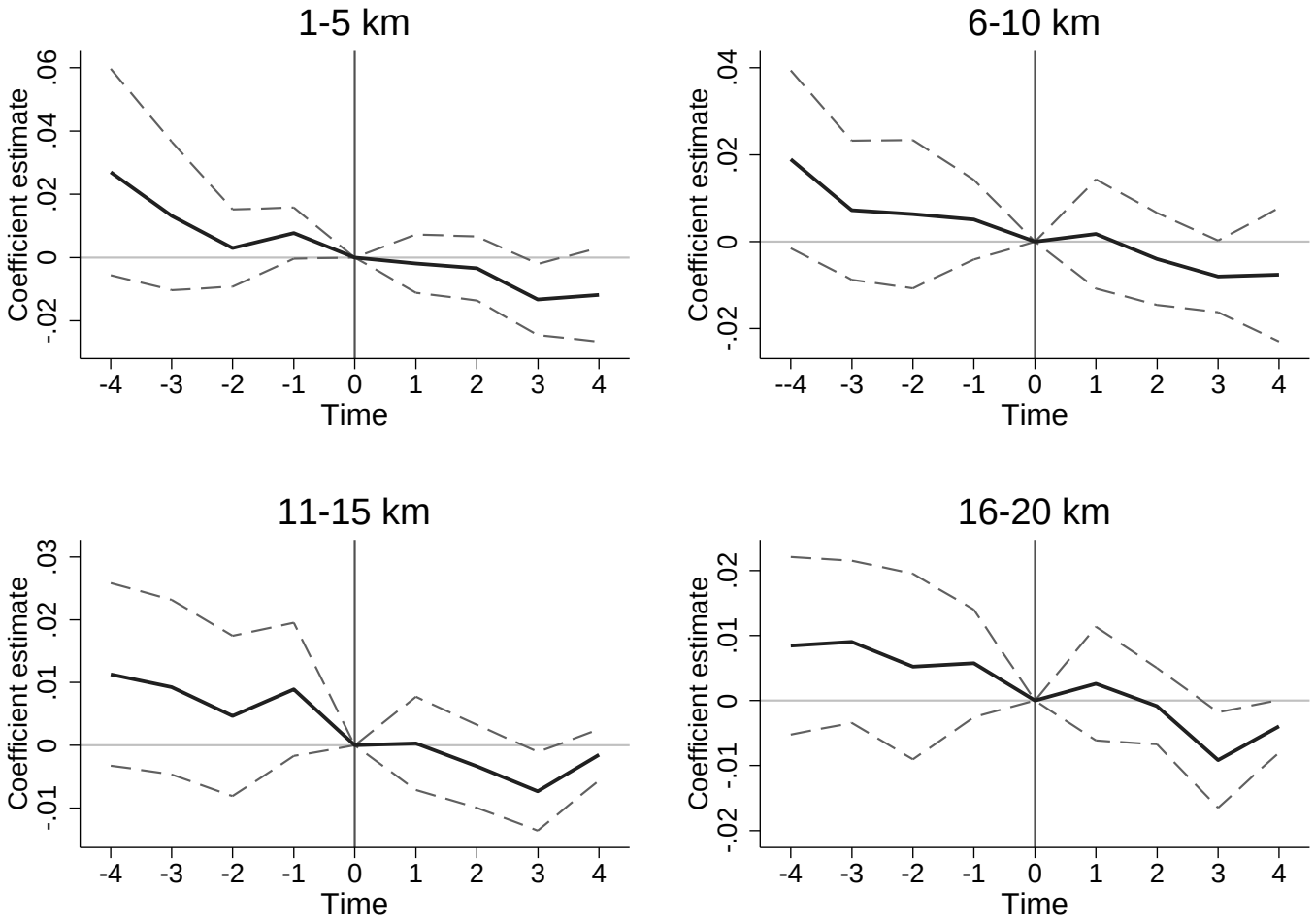
Although I estimated event studies for both specifications described by Equation 3, the figures tended to look very similar across specifications. For this reason, I only report the event study plots for the two-way fixed effects specification here. Additional plots for event study results using linear time trends are available upon request.

Figure A4.1: Coefficient estimates for Equation 3 with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable, tiles in grasslands biome



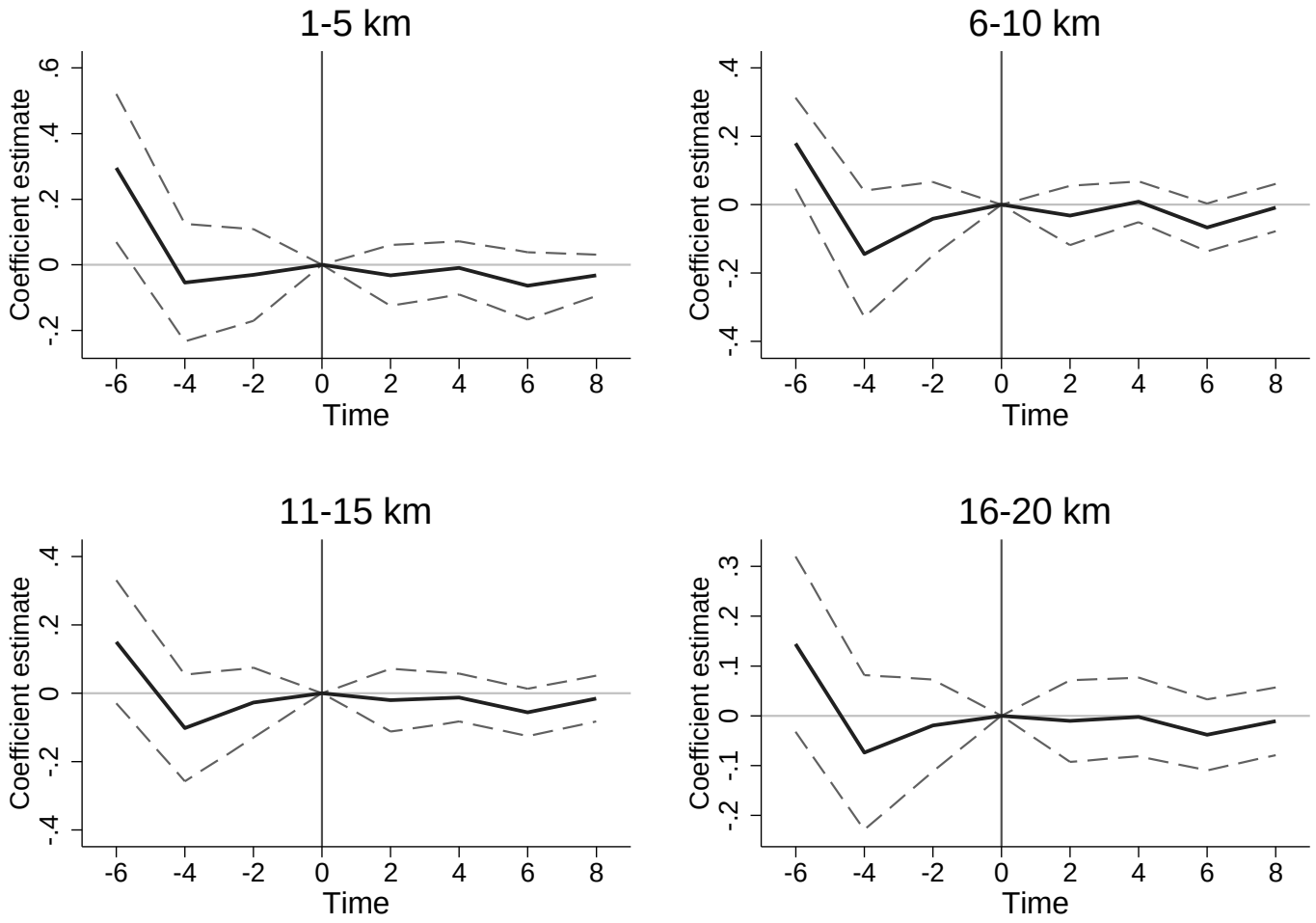
Source: author's calculations using ARD, GFC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. 95% confidence intervals represented by dashed lines. Relative time τ measured in one-year intervals: $\tau = 0$ serves as placebo "before" period.

Figure A4.2: Coefficient estimates for Equation 3 with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable, tiles in rainforest biome



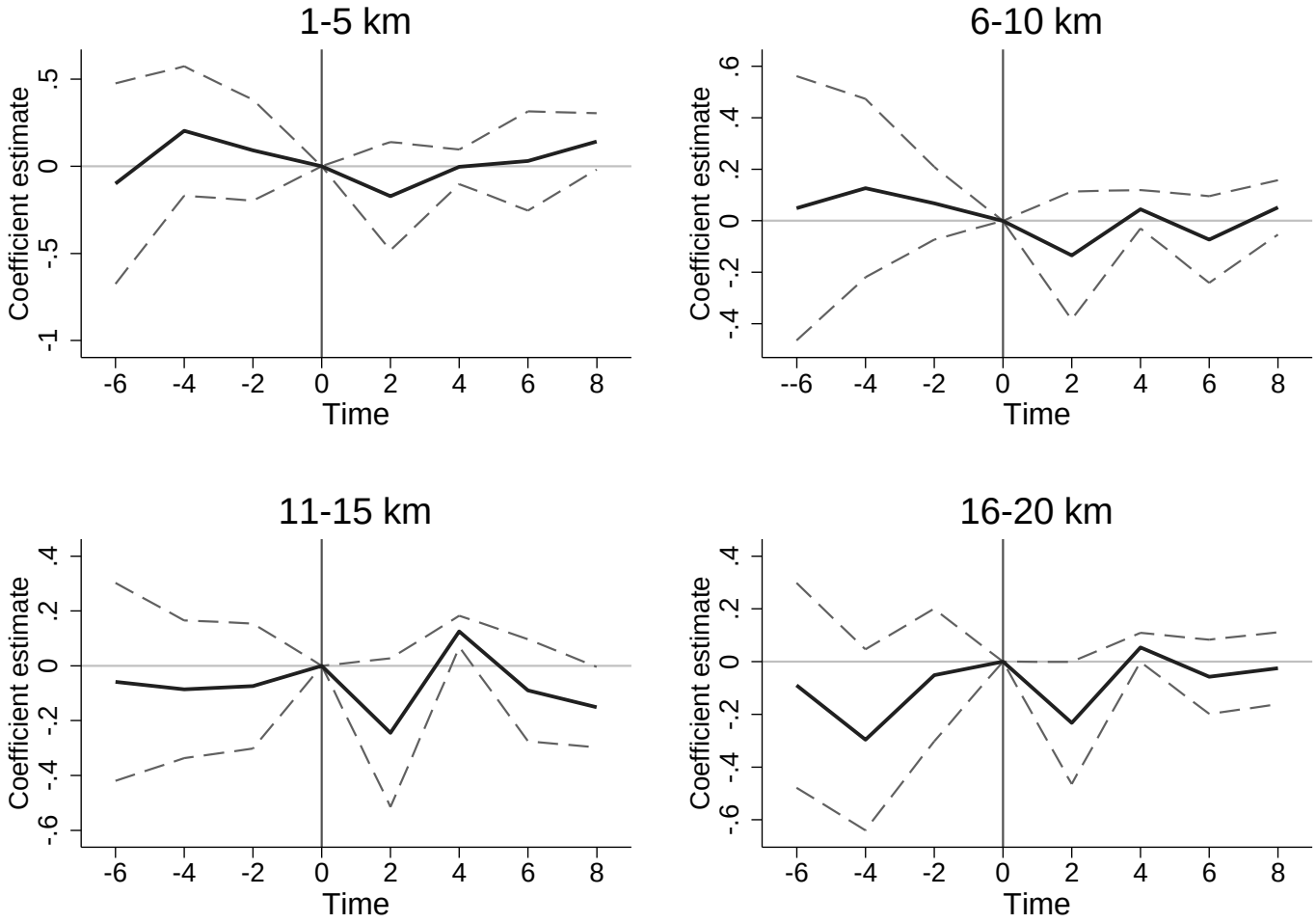
Source: author's calculations using ARD, GFC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. 95% confidence intervals represented by dashed lines. Relative time τ measured in one-year intervals: $\tau = 0$ serves as placebo "before" period.

Figure A4.3: Coefficient estimates for Equation 3 with tile percent forest cover (based on GFCC data) as outcome variable, tiles in grasslands biome



Source: author's calculations using ARD, GFCC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. 95% confidence intervals represented by dashed lines. . Relative time τ measured in two-year intervals: $\tau = -1,0$ serves as placebo "before" period.

Figure A4.4: Coefficient estimates for Equation 3 with tile percent forest cover (based on GFCC data) as outcome variable, tiles in rainforest biome



Source: author's calculations using ARD, GFCC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. 95% confidence intervals represented by dashed lines. Relative time τ measured in two-year intervals: $\tau = -1{,}0$ serves as placebo "before" period.

Appendix A5: Results of robustness test dropping tiles only exposed to camps for two years or fewer

Figure A5.1: Coefficient estimates for Equation 1 with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable, tiles exposed to camps for more than two years

	Grasslands		Rainforests	
	(1)	(2)	(3)	(4)
N camps 1-5 km	0.013	0.012	-0.130	-0.148*
	(0.023)	(0.023)	(0.071)	(0.068)
N camps 6-10 km	-0.001	-0.003	-0.036	-0.055
	(0.009)	(0.009)	(0.058)	(0.054)
N camps 10-15 km	-0.003	-0.004	-0.074	-0.097
	(0.011)	(0.011)	(0.058)	(0.054)
N camps 15-20 km	-0.005	-0.007	-0.099	-0.124*
	(0.009)	(0.009)	(0.054)	(0.052)
Two-way FE?	Yes	No	Yes	No
Linear trend?	No	Yes	No	Yes
Observations	3,521,844	3,521,844	871,632	871,632
R-squared	0.214	0.213	0.218	0.209

Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level and are reported in parentheses. *p<0.05 ** p<0.01 *** p<0.001

Table A5.2: Coefficient estimates for Equation 2 with tile percent forest cover (based on GFCC data) as outcome variable, tiles exposed to camps for more than two years

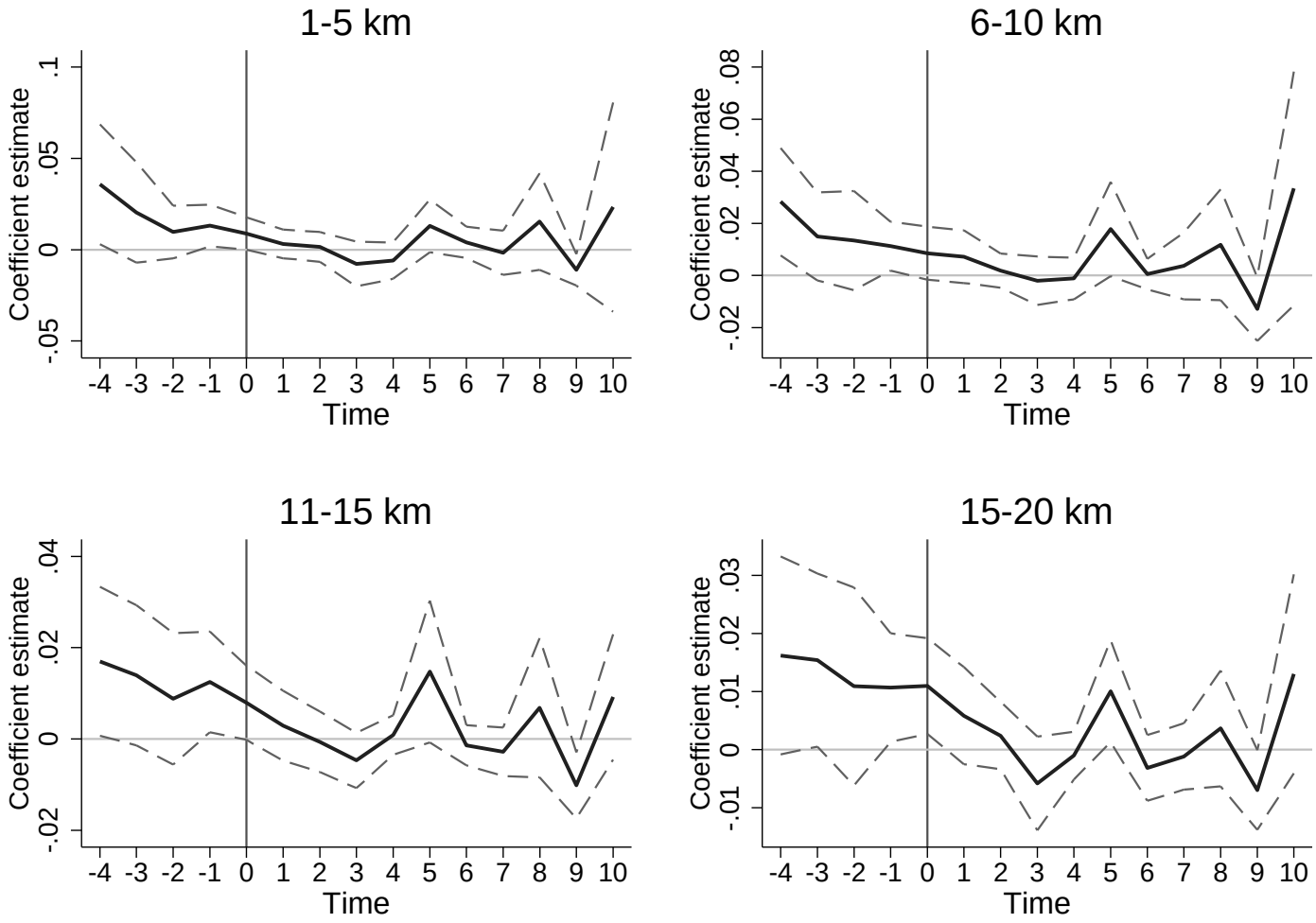
	Grasslands		Rainforests	
	(1)	(2)	(3)	(4)
Yrs. exposure to camp 1-5 km	-0.130*	-0.156**	-0.357*	-0.355*
	(0.052)	(0.054)	(0.148)	(0.159)
Yrs. exposure to camp 6_10 km	-0.101*	-0.128**	-0.315***	-0.313**
	(0.045)	(0.049)	(0.083)	(0.088)
Yrs. exposure to camp 11-15 km	-0.120*	-0.150**	-0.272**	-0.270**
	(0.050)	(0.053)	(0.083)	(0.093)
Yrs. exposure to camp 16-20 km	-0.121*	-0.152**	-0.221*	-0.219*
	(0.053)	(0.056)	(0.090)	(0.090)
Two-way FE?	Yes	No	Yes	No
Linear trend?	No	Yes	No	Yes
Observations	880461	880461	217908	217908
R-squared	0.984	0.984	0.985	0.985

Source: author's calculations based on the ARD, GFCC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level and are reported in parentheses. *p<0.05 ** p<0.01 *** p<0.001

Appendix A6: Additional first difference results

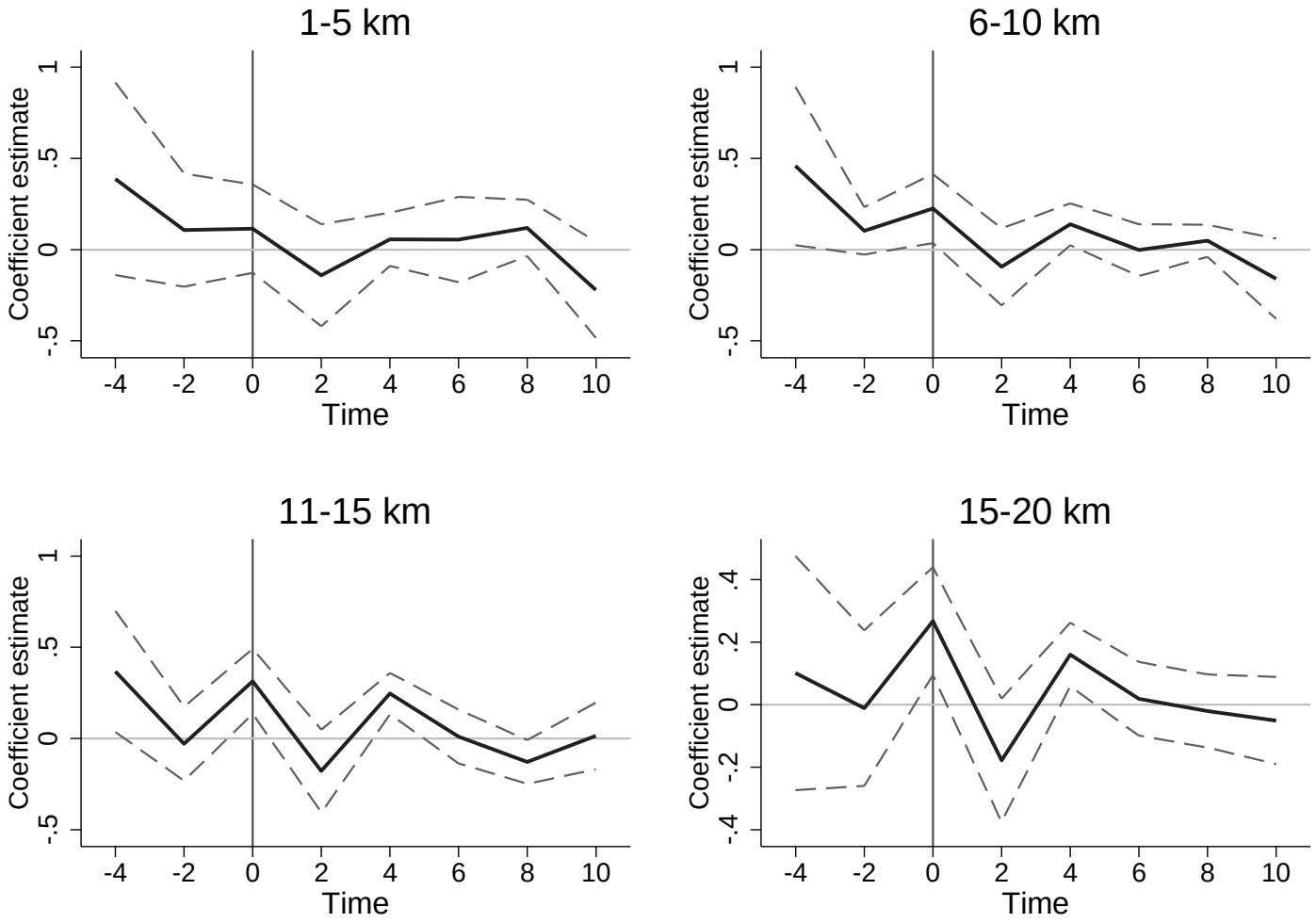
This section provides graphs that illustrate the point estimates and 95% confidence intervals for the first difference specification. This specification takes the difference between exposed and comparison tiles over relative time. It is equivalent to Equation 3 with no placebo "pre-treatment" time period. I report results here for the specification with two-way fixed effects. Outcomes for the specification using a linear trend are extremely similar and are available upon request.

Figure A6.1: Coefficient estimates for first difference specification with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable, tiles in rainforest biome



Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level. 95% confidence intervals displayed with dashed lines.

Figure A6.2: Coefficient estimates for first difference specification with percent forest cover (based on GFCC data) as outcome variable, tiles in rainforest biome



Source: author's calculations based on the ARD, GFCC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level. 95% confidence intervals displayed with dashed lines.

Appendix A7: Results of robustness check estimating outcomes for tiles with ≥ 50 percent forest cover in 2000

In this section, I report results of the main analysis when restricting to sample tiles with 50 percent forest cover or more in 2000 based on the GFC data. The objective of this test is to see whether camp-stimulated changes to the returns to harvesting forest products lead to extraction of high-value areas with high forest density.

Figure A7.1: Coefficient estimates for Equation 1 with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable, tiles with ≥ 50 percent forest cover in 2000

	Grasslands		Rainforests	
	(1)	(2)	(3)	(4)
N camps 1-5 km	-0.060	-0.056	-0.234	-0.247*
	(0.112)	(0.113)	(0.116)	(0.105)
N camps 6-10 km	-0.019	-0.016	-0.091	-0.110
	(0.035)	(0.035)	(0.075)	(0.067)
N camps 10-15 km	-0.019	-0.015	-0.128	-0.147
	(0.056)	(0.057)	(0.084)	(0.073)
N camps 15-20 km	-0.041	-0.035	-0.146	-0.163*
	(0.027)	(0.028)	(0.081)	(0.070)
Two-way FE?	Yes	No	Yes	No
Linear trend?	No	Yes	No	Yes
Observations	582,000	582,000	575,112	575,112
R-squared	0.215	0.213	0.226	0.215

Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level and are reported in parentheses. *p<0.05 ** p<0.01 *** p<0.001

Figure A7.2: Coefficient estimates for Equation 1 with tile percent forest cover (based on GFCC data) as outcome variable, tiles with ≥ 50 percent forest cover in 2000

	Grasslands		Rainforests	
	(1)	(2)	(3)	(4)
Yrs. exposure to camp 1-5 km	-0.368*	-0.434**	-0.454*	-0.431
	(0.148)	(0.139)	(0.205)	(0.216)
Yrs. exposure to camp 6_10 km	-0.150	-0.230*	-0.440***	-0.416***
	(0.098)	(0.097)	(0.100)	(0.113)
Yrs. exposure to camp 11-15 km	-0.115	-0.197*	-0.402**	-0.376*
	(0.086)	(0.088)	(0.113)	(0.141)
Yrs. exposure to camp 16-20 km	-0.104	-0.195**	-0.331**	-0.304*
	(0.075)	(0.072)	(0.108)	(0.125)
Two-way FE?	Yes	No	Yes	No
Linear trend?	No	Yes	No	Yes
GFCC mean 2000	31.78	31.78	51.01	51.01
Observations	145,500	145,500	143,778	143,778
R-squared	0.964	0.963	0.972	0.972

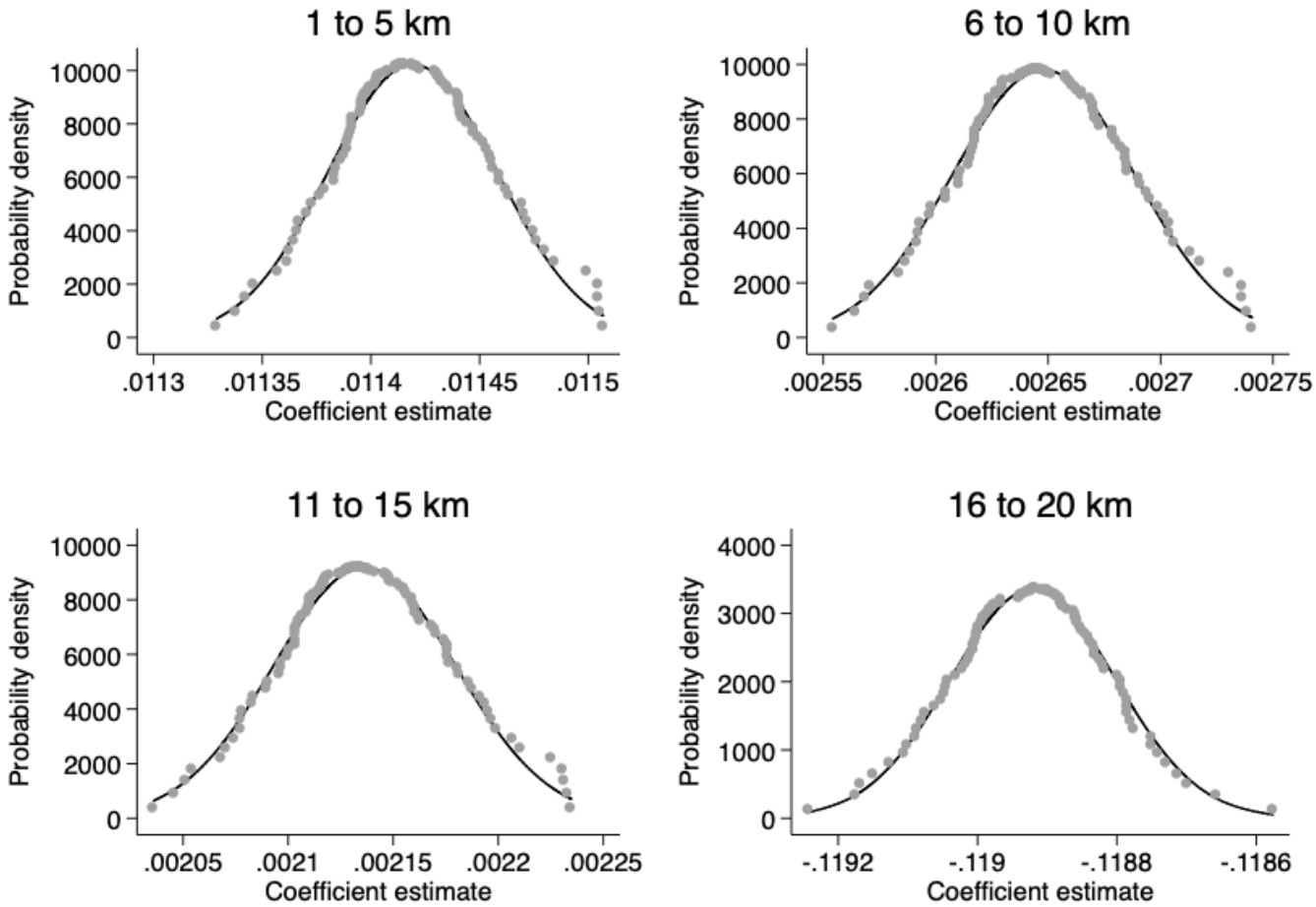
Source: author's calculations based on the ARD, GFCC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level and are reported in parentheses. *p<0.05 ** p<0.01 *** p<0.001

Appendix A8: Results of robustness check using random sample of comparison pixels drawn from outside of camp areas

In Appendix A3 I described my method for generating comparison tiles drawn from buffers with centroids at least 100 km from a camp open 2001-2012 according to ARD. The following steps describe how I used the comparison data to estimate results. First, in my primary dataset of tiles exposed to camps, I dropped all tiles 21-30 km from the nearest camp open at some point 2001-2012. In doing so, I remove the tiles that previously served as the comparison group: this comparison group could be contaminated if some camps are larger than assumed in my identification strategy. Then, I randomly draw 5% of the comparison tiles from the new sample and append these to my primary sample. I estimate Equations 1 and 2 for grasslands and rainforest tiles separately and store the coefficient estimates and standard errors. I repeated this 100 times for each specification. In this section, I plot the distribution of the coefficient estimates for each specification and fit them to a normal distribution.

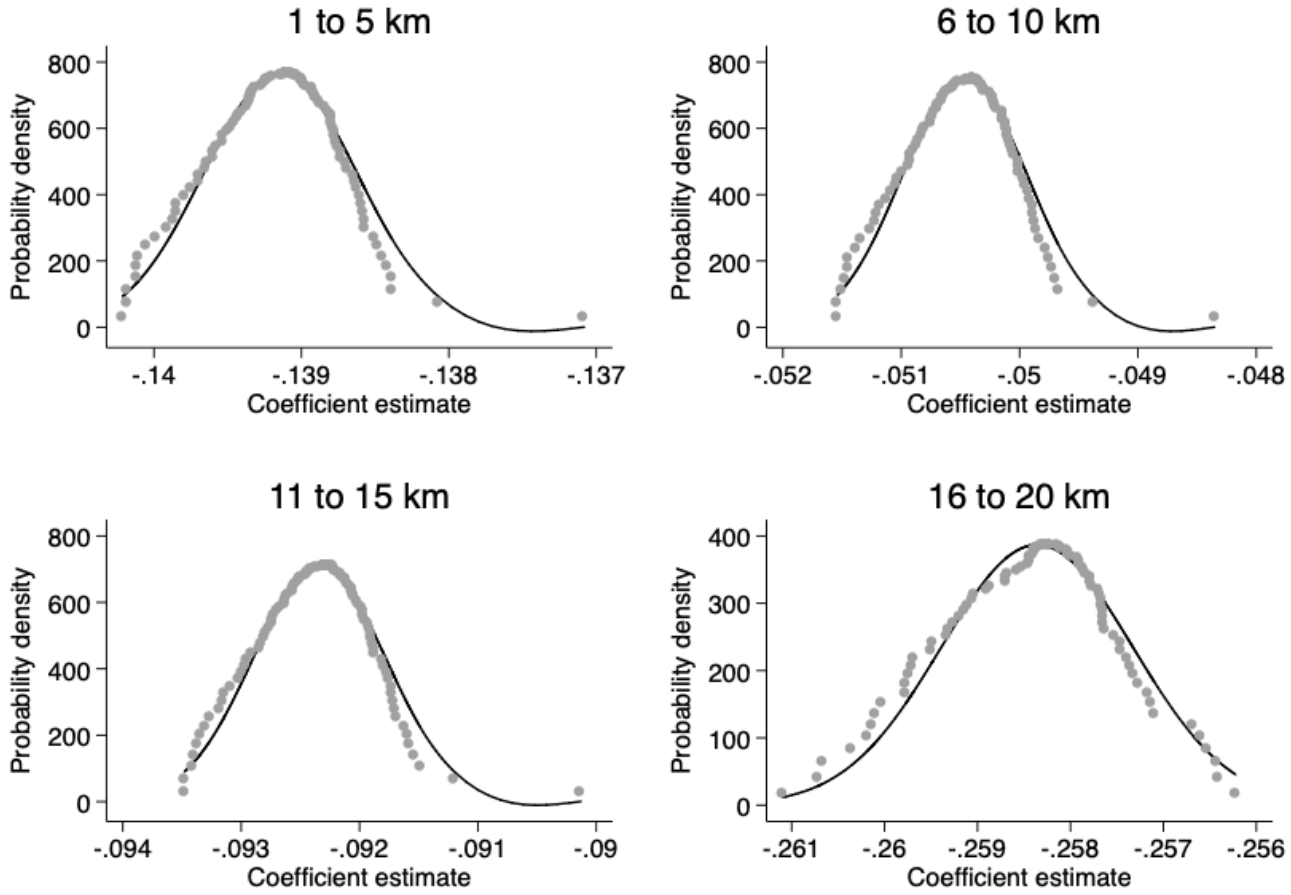
Because the distributions are similar for specifications using two-way fixed effects and a linear time trend, I only report the distribution of coefficients from the specification with two-way fixed effects here. Additional distribution plots available upon request.

Appendix A8.1 Distribution of coefficient estimates for Equation 1 with extensive margin forest loss as the outcome variable (based on GFC data) using comparison tiles from buffers with centroids 100 km from the nearest camp, tiles in grasslands



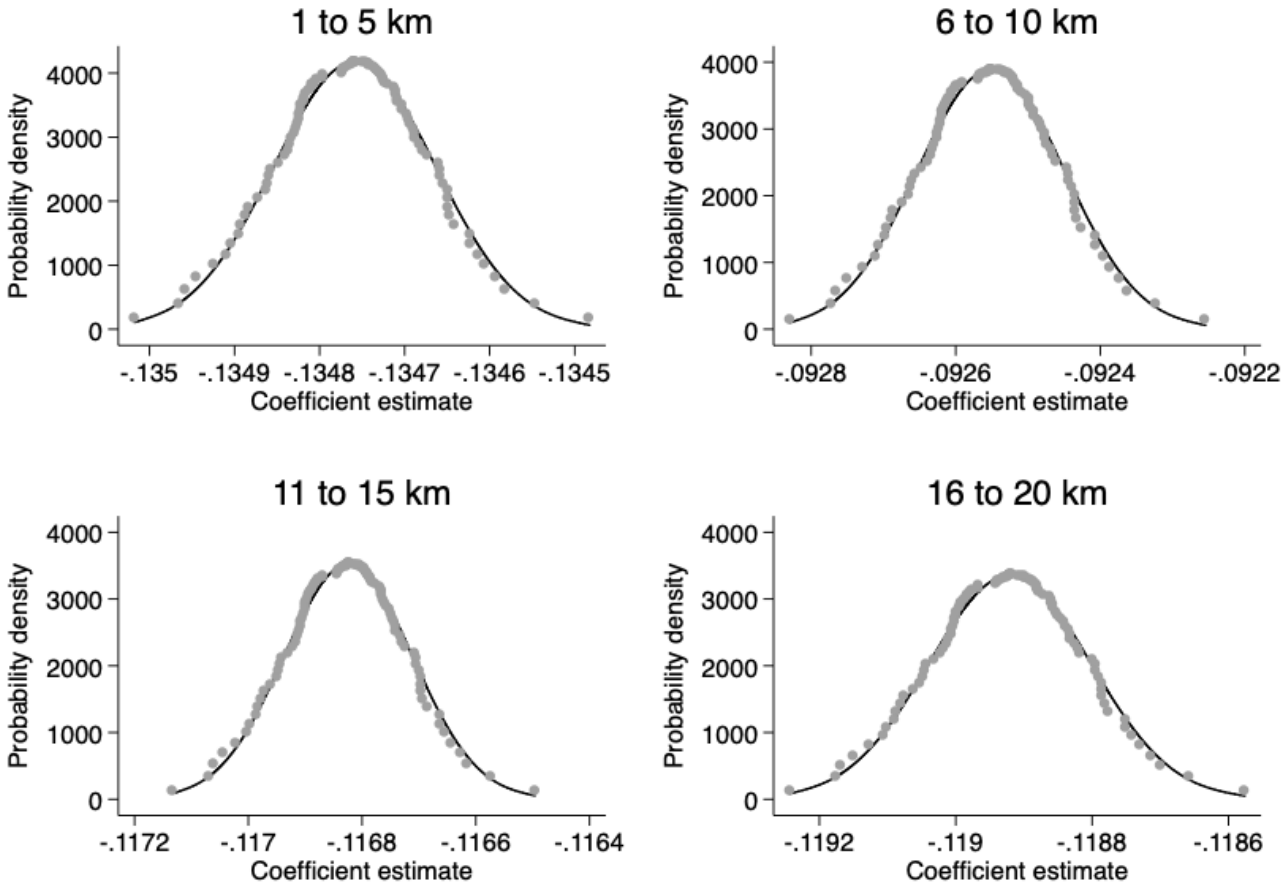
Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use primary sample of tiles measured at 0.01° resolution and drop all tiles 21-30 km from a refugee camp. Comparison tiles are randomly drawn from 30 km buffers with centroids at least 100 km from an ARD camp location. Regression is estimated 100 times to obtain distribution of coefficient estimates. Robust standard errors are clustered at the buffer level.

Appendix A8.2 Distribution of coefficient estimates for Equation 1 with extensive margin forest loss as the outcome variable (based on GFC data) using comparison tiles from buffers with centroids 100 km from the nearest camp, tiles in rainforest



Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use primary sample of tiles measured at 0.01° resolution and drop all tiles 21-30 km from a refugee camp. Comparison tiles are randomly drawn from 30 km buffers with centroids at least 100 km from an ARD camp location. Regression is estimated 100 times to obtain distribution of coefficient estimates. Robust standard errors are clustered at the buffer level.

Appendix A8.3 Distribution of coefficient estimates for Equation 2 with percent forest cover as the outcome variable (based on GFCC data) using comparison tiles from buffers with centroids 100 km from the nearest camp, tiles in grasslands



Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use primary sample of tiles measured at 0.01° resolution and drop all tiles 21-30 km from a refugee camp. Comparison tiles are randomly drawn from 30 km buffers with centroids at least 100 km from an ARD camp location. Regression is estimated 100 times to obtain distribution of coefficient estimates. Robust standard errors are clustered at the buffer level.

Appendix: works cited

Alix-Garcia, Jennifer, Anne Bartlett, and David Saah. 2013. "The Landscape of Conflict: IDPs, Aid and Land-Use Change in Darfur." *Journal of Economic Geography* 13: 589–617.

International Rescue Committee. 2003. "Liberian Orphanages, Handicapped Center Receive Supplies; IRC Aids DP Camps Outside Monrovia." *Reliefweb*.

Kamara, Thomas S. 2003. "Liberia: Civil War and Regional Context."

Marie, Delphine. 2003. "UNHCR Steps up Work around Monrovia before Expanding Inland." *UNHCR*. https://www.unhcr.org/en-us/news/latest/2003/9/3f534d674/unhcr-steps-work-around-monrovia-expanding-inland.html.

news24. 2003. "Rebels Attack Displaced Camp." *News24*.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset." *Journal of Peace Research* 47 (5): 651–660.

UNHCR. 2001. "Liberia in Short."

———. 2003. "Liberia."