H i C N Households in Conflict Network

The Institute of Development Studies - at the University of Sussex - Falmer - Brighton - BN1 9RE

www.hicn.org

Quantifying The Microeconomic Effects of War: How Much Can Panel Data Help?[1]

Margarita Pivovarova[2] and Eik Leong Swee[3]

HiCN Working Paper 116

April 2012

Abstract: The extensive coverage of household surveys in conflict regions in recent decades has fueled a growing literature on the microeconomic effects of war. Most researchers identify these effects using econometric methods, with difference-indifferences – which exploits variation across birth cohorts and war intensity – being the most popular. This paper highlights problems of endogenous war intensity and self-selection due to non-random displacement when using common empirical methods on cross-sectional data, and explains how they can be overcome with panel data. We draw on a unique set of cross-sectional and panel data from Nepal to demonstrate our proposition. Both unobserved locality factors and individual heterogeneity lead to huge swings in the estimates of war intensity effects. Our results imply that researchers ought to think carefully about empirical methods and explain possible statistical biases, especially when their results are used to inform policy decisions. For researchers who use panel data, we propose augmentations to existing methods.

[1] We thank the Central Bureau of Statistics in Nepal for providing us with the survey data, and Gaurab Aryal for supplementary data and suggestions. We also acknowledge comments from Quy-Toan Do, Petros Sekeris, TomWilkening, and participants at Melbourne, Toronto, ADEW (Perth), EEA-ESEM (Oslo), and GECC Conference (Berlin), which improved an earlier version of this paper. All remaining errors are our own.

[2] Corresponding author. Department of Economics, University of Toronto, 150 St. George St., Toronto, Ontario M5S 3G7, Canada. E-mail: rita.pivovarova@utoronto.ca. Tel:+1-4167364217. Fax:+1-4169786713.

[3] Department of Economics, University of Melbourne, Victoria 3010, Australia. E-mail: eswee@unimelb.edu.au. Tel:+61-383445397. Fax:+61-383446899.

1. Introduction

In recent decades, the World Bank conducted household surveys in several war-torn countries around the world, while many more were covered by the Demographic and Health Surveys (Brück, Justino, Verwimp, and Avdeenko, 2010).[1] The extensive coverage of household surveys in conflict regions has fueled a growing literature that addresses many important – but previously unanswered – questions about the microeconomic consequences of war.

To date, more than a dozen studies have examined the effects of violent conflict on the health, schooling attainment, and labour market outcomes of affected populations [see Blattman and Miguel (2010) for a survey of the literature]. Most researchers identify these effects using econometric methods, with difference-in-differences – which exploits variation across birth cohorts and war intensity – being the most popular.[2] In the last five years alone, more than a handful researchers have used this approach, by combining cross-sectional data from household surveys with spatially-varied data on war intensity, to measure the effects on individual outcomes (Merrouche, 2006; Akresh and de Walque, 2008; Bundervoet, Verwimp, and Akresh, 2008; Akbulut-Yuksel, 2009; Swee, 2009; León, 2010; Akresh, Verwimp, and Bundervoet, 2011; Chamarbagwala and Morán, 2011; Shemyakina, 2011; Valente, 2011).

In principle, difference-in-differences accounts for endogeneity that stems from unobserved initial locality factors, which may be correlated with both war intensity and individual outcomes; however, it does not address self selection due to non-random wartime displacement. This issue may be particularly important in the context of war because wartime displacement is widespread and usually unobserved in cross-sectional data. For example, if households that are less able to cope with war tend to be displaced, the proportion of high ability households may be greater

[1]The World Bank's Living Standards Measurement Study covers Azerbaijan, Bosnia & Herzegovina, Guatemala, Iraq, Kosovo, Malawi, Nepal, Peru, Serbia, Tajikistan, and Timor-Leste. Demographic and Health Surveys cover several more, including Burundi, Cambodia, Columbia, Congo, Côte d'Ivoire, Ethiopia, Guatemala, Guinea, Indonesia, Jordan, Liberia, Mali, Mauritania, Mozambique, Peru, Rwanda, Senegal, Sierra Leone, Sri Lanka, Sudan, Tanzania, and Uganda.

[2]While difference-in-differences is commonly used in impact evaluation studies, most researchers rely on variation in policy and time (before and after) until a seminal paper by Duflo (2001) popularised the use of birth cohort variation.

in high intensity regions. This brings about a sample correlation between outcomes and war intensity, which is independent of the effects of war on outcomes (Swee, 2009).

We do two things in this paper. First, we describe the common empirical methods of identifying microeconomic effects of war, explain the challenges, and propose an augmented difference-in-differences specification that addresses lingering identification issues. Second, we draw on cross-sectional and panel data from Nepal to demonstrate that the identification issues are empirically substantial, and can be overcome with the appropriate use of panel data. The data from Nepal is unique in the sense that the cross sections and panel components were collected concurrently within each wave, and that it is one of the few micro panel data sets that covers a war-torn country before and after the onset of war.

Empirically, we find that the correction of endogeneity biases with difference-in-differences is substantial, with corrected estimates taking signs that are opposite to those of naive ordinary least squares. Moreover, when one switches from naive ordinary least squares to difference-in-differences, the magnitude change is more than two-fold. Going from cross-sectional to panel data yields little improvement when one uses standard difference-in-differences; however, the estimate changes significantly when one augments difference-in-differences with individual fixed effects and time-varying location effects, suggesting that self selection may be problematic.

Our results provide two main implications for researchers who are engaged in estimating the microeconomic effects of war. Firstly, we believe that researchers ought to think carefully about empirical methods, especially with regards to what they are identifying and what they can potentially identify. Secondly, we propose that researchers working with cross-sectional data should try to provide ancillary evidence about the extent of possible statistical biases, and interpret estimates appropriately; for researchers who have the option of working with panel data, we propose an augmented difference-in-differences estimator that may be helpful.

The rest of this paper is organised as follows. Section 2 constitutes an introduction of Nepal and the data that we use. Section 3 describes the empirical framework for understanding common empirical methods, and compares them to our preferred specification with panel data. We present the empirical results and conduct robustness checks in Section 4. Section 5 concludes.

2. Background and Data

In this section, we provide a description of Nepal – particularly, of its civil war and public schooling – and the data that we use. Our objective here is to offer sufficient background to understand empirical results, hence the brevity.

2.1 Nepal

Nepal is a small landlocked country in South Asia with a largely agrarian economy. For much of its modern history, Nepal was ruled by a monarchy until widespread protests led to the emergence of multi-party democracy and the introduction of a new constitution in 1990. While democratisation brought expectations of greater political freedom, social mobility, and economic advancement, the new regime faced considerable political instability; there were as many as 12 governments in the first 12 years.

Amidst political turmoil, civil war (otherwise known as the People's War) broke out in 1996, when members of the Communist Party of Nepal (Maoist) attacked a police post in Rolpa district of Western Nepal. In the first few years of civil war, the government mobilised the police to contain the insurgency, but was unable to stop the proliferation of the Maoist propaganda. By 2000, the insurgency covered at least 35 of 75 Nepalese districts. Following King Gyanendra's ascension to the throne in 2001, violence escalated dramatically as the warring parties launched frequent attacks which killed over 4,600 people that year, many of whom were civilians.[3] A ceasefire with the Maoists was reached in 2003, leading to a decline in violence; however, government talks failed and violence resumed. By then, the Maoists were already active in 72 of Nepal's 75 administrative districts.

The next few years saw plenty of violence – in the form of mass strikes, riots, kidnappings, blockades, and terrorist bombings – and gradually, the Maoists dominated the rural regions of Nepal. In September 2005, the Maoists declared a unilateral ceasefire, and began talks with seven major political parties to present a common front against the monarchy. The Nepalese monarch

[3] Informal Sector Service Center Human Rights Yearbook, various issues.

finally relinquished power in April 2006, and a peace agreement was signed, formally ended a decade-long conflict that claimed more than 13,000 lives and displaced thousands more.[4]

Overall, the complexity and length of the war meant that there is considerable variation in war intensity across the country. For instance, there are nearly 5,000 war casualties in the Western region, and only around 1,600 in the Far-Western region. As seen in Figure 1, which depicts the spatial aspect of the conflict intensity by quartiles, the extent of conflict is greater in hilly and rural areas (primarily in Western and Mid-Western Nepal) than in relatively flat and urban ones (in Eastern and Central Nepal).[5]

Where violence was widespread, civilians were compelled to stay home while schools, offices, and factories were shut down. This meant that children who were attending school at the time could be adversely affected, and that those who lived in high intensity regions would have suffered more. The effects of war intensity on schooling attainment could materialise via several channels. Using the categorisation by Swee (2009), these channels could be immediate or aftermath. Immediate channels are typically direct – including the destruction of infrastructure, the displacement of teachers, and a lower demand for schooling – but they could also be indirect, such as war-induced displacement that may lead to disruptions in schooling. Aftermath channels, on the other hand, are associated with cohorts that resume schooling after the war. Given that the later wave of our data is still contemporaneous with war, our results should only reflect immediate channels.

We focus our analysis on public primary and secondary education, which consist of five levels: (i) pre-primary or early childhood education for 3 to 5 years of age; (ii) primary education (Grade 1-5) for children 6 to 10 years; (iii) lower secondary education (Grade 6-8) for eleven to thirteen years-old; (iv) secondary education (Grade 9-10) for fourteen and fifteen year-old; (v) higher secondary education (Grade 11-12) for sixteen and seventeen year-old. Secondary education generally refers to Grades 6-12.[6]

[4]Informal Sector Service Center Human Rights Yearbook, various issues.

[5]See Do and Iyer (2010) on the links between war intensity and regional heterogeneity.

[6]Very few Nepalese attend tertiary education; the average schooling attainment for adults is no more than 4 years

Table 1 provides descriptive statistics comparing trends in primary and secondary schooling outcomes in 1996–2004 across several dimensions. Net enrollment rates and years of schooling in Nepal increased on average during this period; however, children in rural areas are still less likely to attend and complete school, with the difference being more pronounced in secondary school. Moreover, there is substantial variation by region, household per capita expenditure, and war intensity. Notably, while one might expect worse post-war outcomes in high intensity regions, it turns out that is not always the case; for example, we observe higher post-war primary school enrollment rates in high intensity areas. Comparing over time, we also see a larger proportional increase in years of schooling for children in high intensity areas – 48 percent (from 2.21 to 3.28) versus only 14 percent (from 3.49 to 3.97) in low intensity areas – and a reversal of trend in z-scores, with slight improvement in the average z-score in high intensity areas. These statistics suggest that the impact of war intensity on schooling outcomes may not be as adverse as one might expect.

2.2 Data

In this paper, we use data from two main sources. The first comprises two rounds of the Nepalese Living Standards Survey (NLSS); the other is Nepalese conflict data published by the Informal Sector Service Center, a Nepalese non-governmental organisation.

The NLSS is a nationally representative survey of households and communities conducted between June 1995 and June 1996 (NLSS I) and April 2003 and April 2004 (NLSS II) by the Nepal Central Bureau of Statistics with assistance from the World Bank. Importantly, over 91 percent of NLSS I surveys were conducted before the war, making them appropriate baseline observations for identifying effects of the war. The NLSS I and II follow the typical format of other World Bank Living Standards Surveys, and include modules on household composition, income and expenditures, housing, durables, assets, land use and home production. Three types of questionnaires were used: a household questionnaire and two community questionnaires – one for rural areas and another for urban areas. We make use of data on individuals (age, gender, caste, schooling

in 2004.

attainment, and migration history), households (per capita consumption and parental schooling attainment), and communities (local infrastructure and development programs).

The NLSS I contains 3,373 households in 274 sampling units. By design, the NLSS II has two components: first, a nationally representative random cross-section to estimate trends and levels of the socio-economic indicators; this contains 3,912 households from 326 sampling units. Second, a panel sample to track changes in the welfare of the Nepalese population; this contains 1,160 households in 95 sampling units drawn from households in the NLSS I.

The data in this paper consist of all individuals aged 5 and over, with a particular focus on affected cohorts (aged 6-17) whose primary and secondary schooling might have been disrupted by the war.[7] We make use of both cross-sections from the NLSS I and II, and the panel data. There are 15,985 and 17,479 observations in the 1996 and 2004 cross sections respectively (8,494 of whom are in the affected cohorts) and 2,991 individuals in the panel (1,071 of whom are in the affected cohorts). Attrition due to missing covariates is around five percent in both cross sections and 11 percent in the panel. With regards to panel data sampling, around 17 percent of households could not be traced in the second wave. Although replacement households were used to balance household sample size across the two waves, we still lose a significant number of individuals in the panel data as a consequence. We address these issues in Section 4.4.

We use grade-for-age z-score and years of schooling to measure schooling attainment.[8] Z-scores are computed by normalizing the highest grade attended for every individual to his or her gender-cohort reference group, conditional on being in the NLSS I or II. In practice, we prefer the z-score because it collapses information about relative schooling performance for every cohort into a single index (Orazem and King, 2008). On the other hand, years of schooling is (right) censored, so it works better for older cohorts who are unlikely to continue schooling; however, it is a rather noisy measure for younger cohorts. We present results for both measures to be comparable with

[7]Note: the affected cohorts include (i) those aged 6-17 in 1996 (or 14-25 in 2004) because civil war began shortly after the first wave, and (ii) those aged 6-17 in 2004 because the war was ongoing during the second wave. Together, we consider those aged 6-25 in 2004 as affected cohorts.

[8]We do not use enrollment because it is only reported for younger cohorts, thus severely limiting the sample.

the existing literature.

We combine micro data from the NLSS I and NLSS II with district-level data on war intensity from INSEC's Annual Human Rights Yearbooks. Our primary measure of war intensity is constructed by normalising the number of conflict-related deaths from 1996-2006 by the total district population in 1991, the year of the Nepalese census. We also use the number of Maoist abductions in 2002-2006 – again, normalised by the total district population in 1991 – as an alternative measure of war intensity.

Although there are 75 Nepalese districts in total, four districts – Achham, Dolpa, Mustang, and Rasuwa – were not enumerated in the panel.[9] Given our interest in tracking the schooling attainment of individuals who are affected by the war, we need to make sure that the four missing districts do not bias our estimates. We investigate the effect of these missing districts in Section 4.4.

3. Empirical Framework

The identification of microeconomic effects of war poses several challenges. In this section, we describe the problems with using common estimation methods on cross-sectional data, and explain how the appropriate use of panel data can help provide better estimates.

We begin with the premise that one observes a post-war cross-sectional sample of individuals – which is typically the case – and is interested in identifying the effects of war intensity on schooling attainment.[10] The simplest approach may be to estimate the following equation via ordinary least squares (OLS):

$$SCHOOL_{ij} = \beta_1 WAR_j + e_{ij} \tag{1}$$

where $SCHOOL_{ij}$ denotes schooling attainment, WAR_j denotes war intensity, e_{ij} refers to the error term, and i and j are identifiers for individuals and districts respectively. To ensure that the relevant variation in war intensity is used, WAR_j should be constructed to match each individual's

[9]Mustang and Rasuwa are not selected by design (using stratified sampling by geography and ecology) while Achham and Dolpa are missing in 2004 and 1996 respectively, possibly because of the war.

[10]For conciseness, we henceforth omit a vector of (possibly time-varying) individual and household-level controls.

district of residence at the onset of war, even if she was displaced.

Clearly, β_1 identifies war intensity effects only if it is exogenous to unobserved district-level determinants of schooling attainment. This exogeneity assumption is often violated because initial locality conditions, that are correlated with schooling attainment, are also known to be strong determinants of civil conflict (Fearon and Laitin, 2003; Collier, Hoeffler, and Rohner, 2009; Blattman and Miguel, 2010).[11] In Nepal's case, district-level poverty and inequality may be positively correlated with conflict intensity (Murshed and Gates, 2005; Do and Iyer, 2010; Macours, 2011) and negatively correlated with schooling attainment (Pivovarova, 2010), in which case the OLS estimate will be biased downwards.

Given that a random assignment of individuals to districts with varying war intensity is not possible, researchers often adopt one of two common approaches to overcome endogeneity – instrumental variables and difference-in-differences. As it is difficult to come up with novel instruments for war intensity, most studies rely on difference-in-differences (DID).[12] The standard specification exploits variation in birth cohorts – which determines whether individuals are in school at the time of the conflict – and is represented as follows:

$$SCHOOL_{ij} = \beta_2(WAR_j \times AFFECTED_i) + d_j + \epsilon_{ij} \tag{2}$$

where $AFFECTED_i$ is a dummy that equals to one for individuals who are in school at the time of war, d_j denotes district fixed effects, and ϵ_{ij} is the error term. Since district fixed effects are now accounted for, and birth cohorts are arguably exogenous to unobserved determinants of schooling attainment, the DID specification appears innocuous to endogeneity.

Equation (2) represents the quintessential estimating model in the existing literature, but an issue of self selection – due to non-random wartime displacement – remains.[13] This sort of selec-

[11]For instance, assume that $e_{ij} = d'_j + u_{ij}$, where d'_j represents the vector of unobserved district-level determinants. Then, $\text{plim}(\hat{\beta}_1 - \beta_1) = \frac{cov(WAR_j, d'_j)}{var(WAR_j)} \neq 0$ if $cov(WAR_j, d'_j) \neq 0$. We cannot control for d'_j in the OLS because it is multicollinear with WAR_j (or perfectly collinear if WAR_j is binary).

[12]A handful of papers rely on instruments that include spatial variables or lagged policy variables that affect war intensity (Merrouche, 2006; Akresh and de Walque, 2008; Rodriguez and Sanchez, 2009).

[13]If $\epsilon_{ij} = \omega'_i + v_{ij}$, where ω'_i denotes unobserved individual (or household) determinants that are correlated with

tion is often assumed away in difference-in-differences; however, it may be particularly important in the context of war. For example, individuals typically move from high to low intensity municipalities, and households that are better able to cope may have a lower propensity of displacement; thus, the proportion of high ability households may be greater in high intensity districts. Of course, the opposite could also be true if more able individuals tend to have better employment or schooling opportunities outside their municipalities of residence, and are thus more likely to migrate. Either way, this brings about a sample correlation between war intensity and the schooling attainment of affected cohorts, which is independent of the effects of war on schooling attainment (Swee, 2009). As such, the estimation of β_2 is subject to selection bias.

To address selection, the availability of panel data – where individuals are observed before and after the onset of war – is crucial, as one requires multiple observations of the same set of individuals to directly estimate the so-called "unobserved" traits that may influence the propensity of displacement as well as schooling attainment. Of course, if migration history is observed in cross-sectional data, one can provide ancillary evidence on selection by determining whether displacement propensity is influenced by observed characteristics that may also determine schooling attainment. Nevertheless, this approach does not completely rule out selection when pre-war determinants of displacement are unobserved. Therefore, we propose augmenting difference-in-differences with individual fixed effects when panel data are available. The fixed effects (FE) specification is:

$$SCHOOL_{ijt} = \beta_3(WAR_j \times AFFECTED_{it}) + \omega_i + d_{jt} + \varepsilon_{ijt} \tag{3}$$

where ω_i denotes time-invariant individual fixed effects, d_{jt} denotes time-varying district-level effects, and ε_{ijt} is the error term. Given the longitudinal nature of this setup, $AFFECTED_{it}$ is individual and year-specific. All observations prior to the war are assigned $AFFECTED_{it} = 0$; among observations after the onset of war, only those in school are assigned $AFFECTED_{it} = 1$.

Notice that we include the vector d_{jt} because there may well be time-varying district-level

$WAR_j \times AFFECTED_i$, then $\text{plim}(\hat{\beta}_2 - \beta_2) \neq 0$, following similar algebra from the previous footnote.

effects that are relevant, even conditional on individual fixed effects. For instance, post-war reconstruction efforts are often targeted at districts with greater war intensity, and these may have compensating effects on the post-war outcomes of affected cohorts residing in high intensity districts (Miguel and Roland, 2011). Alternatively, returns to education may differ across high and low intensity districts (especially after the onset of war), and these differences may be captured by time-varying district-level effects.

Equation (3) is our preferred specification for identifying microeconomic effects of war as it offers a strategy to overcome endogeneity as well as other lingering identification issues.

4. Empirical Estimation

We proceed to carry out the empirical estimation. All regressions include individual and household-level controls such as sex, parental education level, per capita consumption, and caste. Cohort and year fixed effects are also accounted for when appropriate.

4.1 OLS and DID

We begin with the regressions in which years of schooling is the measure of schooling attainment (left panel of Table 2). We replicate common estimation methods – OLS and DID – by using the cross-sectional data in the NLSS. Focusing only on the 2004 cross section, we find that the OLS estimate is -0.227, statistically significant at one percent level [column (1)]. While this might hint towards negative effects of war intensity on schooling attainment, we know that the estimate is subjected to endogeneity bias in the sense that unobserved locality factors may be driving both war intensity and years of schooling. Indeed, when we use DID [column (2)], the effect turns positive (0.278, statistically significant at five percent level). To increase statistical power, we repeat this exercise by pooling the 1996 and 2004 cross sections and arrive at similar conclusions [columns (3) and (4)]. Together, these results confirm the presence of a negative endogeneity bias.[14]

[14]We run further tests for the presence of endogeneity, by estimating a system of seemingly unrelated regressions (SUR) with war intensity and grade-for-age z-score as dependent variables, and district-level covariates presumably correlated with them for both survey years. Although separate OLS regressions yield coefficients which are almost identical to SUR estimates, we find that the error terms in the two regressions are not independent (confirmed by the Breusch Pagan test), implying the presence of unobserved factors which influence both war intensity and schooling

Even though it is unsurprising to find negative endogeneity bias, the one we have at hand is rather large in comparison to the existing literature, as most studies still find significant negative effects after correcting for endogeneity [see for example, Akresh and de Walque (2008), Bundervoet, Verwimp, and Akresh (2008), Swee (2009), Akresh, Verwimp, and Bundervoet (2011), Chamarbagwala and Morán (2011), and Shemyakina (2011)]. To illustrate the extent of this bias, consider the following: the effect of a one standard deviation increase in war casualty – around 0.700 deaths per thousand – on years of schooling changes from -0.156 (OLS) to 0.191 (DID), which is more than a two-fold difference [columns (1)-(2)].[15]

Next, we use the second wave of the NLSS panel data (effectively treating it as a cross section) and reestimate OLS and DID [columns (1)-(2) of Table 3]. Again, we find that the OLS estimate is negative (-0.260, statistically insignificant) while the DID estimate is positive (0.782, statistically significant at the five percent level). Both of these are larger in magnitude than their cross-sectional counterparts, and the reason for this is unclear. One possibility is that survey quality is lower in the panel sample, as it is difficult to track down panel households during the war. However, according to Hatlebakk (2007), data quality is consistent across NLSS I and II, and across Maoist and non-Maoist districts, so this might not be too problematic. Alternatively, differences between the cross-section and panel estimates may be a result of attrition bias, as households from the top of the income distribution may be more likely to migrate overseas. Nevertheless, we do not find convincing evidence in favour of this explanation.[16]

Moving on, we repeat the estimation by using grade-for-age z-scores as the measure of school-

attainment (Appendix Table 1).

[15]It is plausible that the positive DID result may be partly driven by the Maoist agenda of providing universal access to schooling, particularly for the lower castes and girls. However, Valente (2011) uses repeated cross-sectional data from the Demographic and Health Surveys, and finds positive war intensity effects on female schooling attainment but not enrollment; she argues, thus, that the Maoist agenda was not the key factor in determining war intensity effects.

[16]We run attrition propensity regressions, to compare households that are observed in both waves of the panel and those that are only observed in the first wave (attrited). We find no correlation between attrition and pre-determined attributes – such as parental schooling attainment, per capita consumption, caste, and landholding – controlling for district fixed effects.

ing attainment (right panel of Table 2). Using cross-sectional data, we find that the OLS estimates are again negative (between -0.056 and -0.065, depending on sample), while the DID estimates are positive (between 0.060 and 0.058, depending on sample). Moreover, as z-scores are constructed from schooling attainment, of which one standard deviation is approximately four years of schooling, the magnitudes of these estimates are equivalent to 0.22-0.26 years of schooling, which are consistent with those in the left panel of Table 2. Reestimating the OLS and DID using only the first wave of the panel yields similar conclusions as before [columns (5)-(6) of Table 3].

4.2 Individual Fixed Effects

Next, we run our preferred FE specification using both waves of the NLSS panel data. Again, we examine both years of schooling and grade-for-age z-scores (see Table 3). Accounting for individual fixed effects, the estimate in column (3) now tends towards zero (0.061, statistically indistinguishable from zero). This implies that there are no significant war intensity effects on years of schooling, after correcting for selection bias. This conclusion is robust to the inclusion of time-varying district-level effects, as the FE estimate in column (4) continues to be statistically indistinguishable from zero.[17] Turning over to columns (7)-(8), we see that there are no significant effects of war intensity on grade-for-age z-score either.

Overall, FE estimates tend towards zero as compared to their DID counterparts. These results imply that the selection bias is positive, and is possibly driven by the self selection of less able households out of high intensity districts into low intensity districts. If one believes that high ability households have better distant networks that provide better migration opportunities or lowing moving costs (Chiquiar and Hanson, 2005; McKenzie and Rapoport, 2010), then this particular direction of selection may not make sense; however, as the Nepalese economy is overwhelmingly agrarian, many of these high ability households may have immobile farm assets that could poten-

[17] In fact, the FE estimate increases only slightly in magnitude to 0.068, suggesting that time-varying district-level factors are unimportant at best. This is perhaps unsurprising since the later wave of the data is still contemporaneous with civil war, so reconstruction projects may not have kicked in yet. In addition, returns to education may not have decreased in high intensity districts as the Maoists were known for their pro-education stance.

tially make out-migration very costly.[18]

To further examine the issue of self selection, we consider wartime displacement as a function of individual ability and war intensity in the origin district. The model we have in mind decomposes ability into (i) skills, that manifest through higher parental schooling or high-paying occupations, and (ii) landholding, which signals the accumulation of wealth for able individuals. Using cross-sectional data that contains individual-level information on displacement during the war, we run a series of displacement propensity regressions (shown in Table 4). Accounting for district fixed effects and other individual and household-level covariates, we find that skills are positively correlated with displacement while landholding is negatively correlated with displacement [columns (1)-(3)]. These results suggest that skills and landholding have contrasting relationships with displacement in general. Moreover, when we interact skills and landholding with war intensity, the first-order correlations diminish; instead, skills and landholding influence displacement propensity depending on war intensity [columns (4)-(6)]. In particular, land owners are less likely to be displaced regardless of skills, which is consistent with the fact that costly out-migration may be a determining factor in finding a positive selection bias.[19]

If it is true that the selection bias is driven by the displacement story that we proposed above, then we should be able to find some critical level of ability above which a sample of so-called "high ability" individuals would exhibit no selection bias. To explore this idea, we run a series of DID regressions using the second wave of the NLSS panel data, and present the war intensity estimates

[18]We acknowledge that there could be another form of self selection which is driven by individuals moving overseas during the war (thus being dropped from our sample). For example, individuals at the bottom end of the ability distribution may migrate overseas in response to war, and this will result in an over-representation of high ability individuals in high intensity districts, which in turn causes positive selection bias. That said, we think this is an unlikely scenario given that (i) high ability individuals are probably more able to pay the high costs of overseas migration, and (ii) attrition in the panel data is uncorrelated with parental schooling or landholding, implying that those attrited from the panel may not be that different in terms of ability (see Section 4.4). We thank Pushkar Maitra for raising this point.

[19]Notice that these results are quite robust to different measures of skills, whether pre-determined (parental secondary schooling) or not (skilled and non-agricultural occupation). This alleviates somewhat the concern that the skill measures observed in 2004 may be endogenous to war intensity.

on years of schooling and grade-for-age z-scores in Figure 2. We start with all individuals and then sequentially remove individuals from the lower end of the ability distribution as measured by years of parental schooling.[20] We find that war intensity effects are rather stable when we remove individuals whose parents have up to five years of schooling. Beyond five years of parental schooling, however, the DID coefficients become statistically insignificant. Strikingly enough, we obtain the same critical point of five years regardless of the measure of schooling attainment. This finding provides support to the model of self-selection when less able individuals move out of high intensity districts thus generating a positive selection bias in the DID regressions.

While unpacking the mechanisms behind war effects is not our primary objective in this paper, we are interested in whether specific aspects of the war may have cushioned the war intensity effects on schooling attainment. In particular, the Maoists were known to be strong advocates for social change and equal access to education, and this could have counteracted the negative impacts of war intensity. To investigate this hypothesis, we follow Valente (2011) by replacing the number of war casualties with the number of Maoist abductions in 2002-2006. Both measures are observed at the district-level and are normalised by the total district population in 1991. Table 5 presents the OLS, DID, and FE results using the format of Table 3. In general, the estimates suggest that war intensity – as proxied by Maoist abductions – has no distinguishable effect on schooling attainment; however, it appears that grade-for-age z-scores may actually have slightly improved in districts with more Maoist abductions. We believe, therefore, that the specific nature of rebel agenda in this case may explain why we do not find negative war intensity effects in Nepal.[21]

[20] The first coefficient in each graph corresponds to the DID estimate in columns (2) and (6) of Table 3 respectively. The sample size of 2,991 then gradually declines to 192; that is, only 192 individuals have parents with at least 11 years of schooling. Beyond that, the sample becomes too small to estimate the DID model.

[21] Another explanation for the absence of war intensity effects could be due to some sort of convergence in schooling attainment across high and low intensity districts. To investigate this, we run a set of falsification tests in Appendix Table 3. If indeed certain districts are catching up in schooling attainment (in the absence of war) and that confounds our estimates of war intensity effects, we may be able to detect this via trends in schooling attainment. Our approach here is to consider rural versus urban districts, in place of high versus low intensity districts. The FE estimates [columns (3)-(4)] suggest that rural districts are actually doing worse in years of schooling after the onset of war, which signals to

In summary, we have shown here that endogeneity and self selection issues form substantial barriers towards empirical identification. OLS estimates are negative while DID estimates are positive, confirming a negative endogeneity bias. The FE estimates, on the other hand, are statistically indistinguishable from zero, suggesting that the selection bias is positive.[22]

There are two main lessons so far. Firstly, one should be very careful about estimation biases in examining war intensity effects as we have clearly demonstrated that such biases could lead to huge swings in policy-relevant estimates. Secondly, the FE estimations reveal that wartime displacement depends crucially upon unobserved ability, suggesting that individuals moved because of economic opportunities rather out of fear. This makes sense because the People's War was a low-casualty episode of ideological conflict, rather than a full-blown civil war, in which case one may not have found such heterogeneous response to war intensity.[23]

4.3 Cohort and Gender Analyses

While the FE estimates suggest that there are no statistically significant effects of war intensity on schooling attainment, it may be worthwhile to examine the effects by cohort and gender, in case cohort or gender-specific results are hidden beneath the average effect.

First, we decompose the effects by birth cohort. We replace the dummy for affected cohorts $AFFECTED_{it}$ by a set of cohort dummies for those aged 6-17 in 1996 (or 14-25 in 2004) and those aged 6-17 in 2004. The complete set of affected cohorts thus include 20 birth cohorts (aged 6-25 in 2004), as represented by 20 cohort dummies $\{AFFECTED_{it}^{1}, \cdots, AFFECTED_{it}^{20}\}$. Using both

us that the convergence story is perhaps untrue.

[22]Note: Although we find evidence of low ability individuals "self-selecting" out of high intensity districts, this does not appear to cause wealth equalisation across high and low intensity districts (Gini coefficients in assets do not converge over time).

[23]In addition, we find that the heterogeneous displacement response to war intensity is unaffected by insurgent influence, which further confirms that individuals did not move out of fear (see Appendix Table 3). Here, we estimate displacement propensity by limiting the sample to (i) Maoist-controlled districts [as defined in Hatlebakk (2007)], and (ii) districts in which the United People's Front (UPF) – militant arm of the Communist Party of Nepal – held constituent seats in 1991 [using data from Whelpton (2005)]. None of these samples deliver the sort of selection by ability that we found in Table 4.

waves of the NLSS panel data, we then run the FE regression below, focusing on the interaction of these cohort dummies with war intensity:

$$SCHOOL_{ijt} = \sum_{1}^{20} \beta_3^k (WAR_j \times AFFECTED_{it}^k) + \omega_i + d_{jt} + \varepsilon_{ijt} \tag{4}$$

Column (1) of Table 6 presents the FE results based on the full panel data, with years of schooling being the dependent variable. By and large, the results suggest that there are no significant cohort-specific war intensity effects, regardless of cohort. These FE estimates suggest that results found previously in Table 3 are not a facade of averaging over cohorts that may have been positively or negative affected by war intensity. As a quick falsification check, we also include 5 more cohort dummies (for those aged 26-30 in 2004) in our analyses. These cohorts would have been too old to experience war intensity effects on schooling attainment, and the results confirm this. Next, we repeat the above exercise by splitting the sample by gender. Columns (2) and (3) presents the estimates for the male and female sample respectively. Here, the estimates of cohort-specific war intensity effects for males approximate those for the full sample.

Turning to grade-for-age z-scores as the dependent variable, we reach similar conclusions about cohort-specific war intensity effects, except that the 16-year-old males experience a marginally significant negative effect of -0.459 [columns (5) of Table 6]. Interestingly, this male cohort would have been 8-years-old when the civil war began in 1996, which means that they were exposed to war for the entire period of their primary schooling lives. None of the females cohorts display a similar pattern.

4.4 Robustness

In this brief section, we carry out a number of robustness checks to see if the results of our estimation are sensitive to the definition of affected cohort, attrition due to missing covariates, unbalanced panel, or non-enumerated districts. The main results are shown in Table 7, where column (1) represents our earlier FE results for easy comparison. For conciseness, we focus on the effects of war intensity on grade-for-age z-score.

First, we change the definition of the affected cohort by including 18 year-old [column (2)] and 19 year-old [column (3)] individuals, to account for possible late exit from school. We cannot examine early entry because the second wave of the panel only includes individuals aged seven and above. Estimation results are robust to the deviation from our original definition: we find no direct effect of war intensity on grade-for-age z-score, even after including the older cohorts, as implied by FE estimates.

In column (4), we run the FE regression but drop observations from the capital district of Kathmandu. There are several reasons why one might check robustness by doing this. First, Kathmandu is the effectively the central command in the fight against the insurgents, which makes it unique important in the context of this war. Second, it is the most urbanized district in our sample with one of the lowest war intensity, so it might attract large numbers of displaced persons during the war. We are thus concerned that our estimates may be confounded due to any of these complex relationships, but the results in column (4) suggest that dropping Kathmandu makes very little difference to our conclusion.

Having dropped a fair number of observations missing one or more covariates, we acknowledge that this may bias our results. In response, we regress a dummy variable that denotes having missing covariates on the standard set of controls. The FE results in column (5) demonstrate that the relationship between the incidence of missing covariates and the interaction of war intensity and affected cohort is negative but statistically insignificant. This implies that, conditional on being in the affected cohorts, individuals from low intensity districts are no more likely to report missing covariates.

Next, we test for the presence of attrition bias due to an unbalanced panel. We do this by estimating the relationship between attrition from the panel and the interaction of war intensity and affected cohort [column (6)]. The FE estimate supports the hypotheses of random exit from the panel survey. In addition, we find no correlation between attrition and predetermined characteristics such as parental schooling attainment and landholding. We conclude that the effects found in our main specification are unlikely to be influenced by the unbalanced panel observations.

Lastly, we address the issue of missing districts in the sample. Our main concern is that omitted observations from the districts of Achham and Dolpa may be empirically significant as these are high intensity districts (with normalised war causalities of approximately 1.719 and 2.419 deaths per thousand respectively). We construct dummy variables denoting observations from Acham (in 1996) and Dolpa (in 2004), and regress them on the standard set of controls, to test for the correlation between residence in Achham and Dolpa, and the interaction of war intensity and affected cohort Although statistically significant at the 10 percent level, the negative correlation between residence in Achham and the interaction term is rather small [column (7)]. There seem to be no association between residence in Dolpa and the interaction term [column (8)].

5. Conclusions

In this paper, we show that identification issues, such as endogeneity and self selection, are important in estimating microeconomic effects of war. The standard difference-in-differences estimates account for the endogeneity of war intensity to unobserved locality factors, but they do not address the issue of self selection due to non-random wartime displacement. In particular, we find a significant positive selection bias in the standard difference-in-differences estimates, which suggests that unobserved individual heterogeneity is important, and that those at the lower end of the ability distribution are more likely to be displaced. This suggests that low ability individuals are "twice-cursed" because they not only experience direct effects but are also more likely to incur indirect (adjustment) costs of displacement. This may be specific to Nepal because the People's War was a low-casualty episode of ideological conflict, rather than a full-blown civil war, in which case one may not find such heterogeneous response to war intensity.

With the emergence of more micro panel data sets among war-torn countries in recent years, the execution of relevant empirical methods will become increasingly important. Researchers who have panel data may want to consider augmenting difference-in-differences with individual fixed effects to address self selection. Others who work with cross-sectional data ought to think carefully about empirical methods, especially with regards to what they are identifying and what they can potentially identify. They should also try to provide ancillary evidence about the extent of statistical biases, and provide accurate interpretations of empirical estimates, especially when these are used to inform policy decisions.

References

AKBULUT-YUKSEL, M. (2009): "Children of War: The Long-Run Effects of Large-Scale Physical Destruction and Warfare on Children," *Households in Conflict Network Working Paper 62*.

AKRESH, R., AND D. DE WALQUE (2008): "Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide," *World Bank Policy Research Working Paper No. 4606*.

AKRESH, R., P. VERWIMP, AND T. BUNDERVOET (2011): "Civil War, Crop Failure, and Child Stunting in Rwanda," *Economic Development and Cultural Change*, 59(4), 777–810.

BLATTMAN, C., AND E. MIGUEL (2010): "Civil War," *Journal of Economic Literature*, 48(1), 3–57.

BRÜCK, T., P. JUSTINO, P. VERWIMP, AND A. AVDEENKO (2010): "Identifying Conflict and Violence in Micro-Level Surveys," *Households in Conflict Network Working Paper 79*.

BUNDERVOET, T., P. VERWIMP, AND R. AKRESH (2008): "Health and Civil War in Rural Burundi," *Journal of Human Resources*, 44(2), 536–563.

CHAMARBAGWALA, R., AND H. E. MORÁN (2011): "The Human Capital Consequences of Civil War: Evidence from Guatemala," *Journal of Development Economics*, 94(1), 41–61.

CHIQUIAR, D., AND G. H. HANSON (2005): "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113(2), 239–281.

COLLIER, P., A. HOEFFLER, AND D. ROHNER (2009): "Beyond Greed and Grievance: Feasibility and Civil War," *Oxford Economic Papers*, 61(1), 1–27.

DO, Q.-T., AND L. IYER (2010): "Geography, Poverty and Conflict in Nepal," *Journal of Peace Research*, 47(6), 735–748.

DUFLO, E. (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4), 795–813.

FEARON, J. D., AND D. D. LAITIN (2003): “Ethnicity, Insurgency, and Civil War,” *American Political Science Review*, 97(1), 75–90.

HATLEBAKK, M. (2007): “LSMS Data Quality in Maoist Influenced Areas of Nepal,” *Chr. Michelsen Institute Working Paper 2007:6*.

INFORMAL SECTOR SERVICE CENTER (various issues): *Human Rights Yearbook*. Kathmandu, Nepal.

LEÓN, G. (2010): “Civil Conflict and Human Capital Accumulation: The Long Term Effects of Political Violence in Perú,” *Journal of Human Resources*, forthcoming.

MACOURS, K. (2011): “Increasing Inequality and Civil Conflict in Nepal,” *Oxford Economic Papers*, 63(1), 1–26.

MCKENZIE, D., AND H. RAPOPORT (2010): “Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks,” *Review of Economics and Statistics*, 92(4), 811–821.

MERROUCHE, O. (2006): “The Human Capital Cost of Landmine Contamination in Cambodia,” *Households in Conflict Network Working Paper 25*.

MIGUEL, E., AND G. ROLAND (2011): “The Long Run Impact of Bombing Vietnam,” *Journal of Development Economics*, 96(1), 1–15.

MURSHED, S. M., AND S. GATES (2005): “Spatial Horizontal Inequality and the Maoist Insurgency in Nepal,” *Review of Development Studies*, 9(1), 121–134.

ORAZEM, P. F., AND E. M. KING (2008): “Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy,” in *Handbook of Development Economics*, ed. by T. P. Schultz, and J. A. Strauss, vol. 4, chap. 55, pp. 3475–3359. Elsevier.

PIVOVAROVA, M. (2010): “Caste, Gender and School Enrollment: Evidence from the Nepalese Living Standard Survey,” *University of Toronto mimeo*.

RODRIGUEZ, C., AND F. SANCHEZ (2009): "Armed Conflict Exposure, Human Capital Investments and Child Labor: Evidence from Colombia," *Households in Conflict Network Working Paper 68*.

SHEMYAKINA, O. (2011): "The Effect of Armed Conflict on Accumulation of Schooling: Results from Tajikistan," *Journal of Development Economics*, 95(2), 186–200.

SWEE, E. L. (2009): "On War and Schooling Attainment: The Case of Bosnia and Herzegovina," *Households in Conflict Network Working Paper 57*.

VALENTE, C. (2011): "What Did the Maoists Ever Do for Us? Education and Marriage of Women Exposed to Civil Conflict in Nepal," *The University of Sheffield, Department of Economics Working Paper 2011009*.

WHELPTON, J. (2005): *A History of Nepal*. Cambridge University Press, New York.

Table 1 - Descriptive Statistics on Schooling Outcome

Categories	Net enrollment rates				Years of schooling	Grade-for-age z-score
	Primary	Lower secondary	Secondary	Higher secondary		
Total	0.75 (0.59)	0.32 (0.23)	0.17 (0.13)	0.11 (0.04)	3.73 (2.85)	0.00 (0.00)
Gender:						
Male	0.80 (0.68)	0.34 (0.27)	0.19 (0.16)	0.12 (0.04)	3.82 (3.10)	0.00 (0.00)
Female	0.71 (0.50)	0.29 (0.17)	0.16 (0.10)	0.09 (0.04)	3.22 (2.26)	0.00 (0.00)
Urban	0.83 (0.75)	0.48 (0.41)	0.34 (0.28)	0.25 (0.14)	4.61 (4.09)	0.39 (0.53)
Rural	0.73 (0.56)	0.26 (0.19)	0.12 (0.08)	0.05 (0.01)	3.20 (2.37)	-0.12 (-0.12)
Region:						
Eastern	0.77 (0.60)	0.31 (0.27)	0.19 (0.18)	0.06 (0.01)	3.56 (2.84)	0.03 (0.05)
Central	0.68 (0.58)	0.35 (0.26)	0.17 (0.17)	0.15 (0.07)	3.33 (2.85)	-0.10 (0.03)
Western	0.85 (0.71)	0.34 (0.23)	0.17 (0.13)	0.13 (0.03)	4.06 (3.09)	0.24 (0.24)
Mid-Western	0.82 (0.52)	0.22 (0.17)	0.15 (0.02)	0.03 (0.01)	3.27 (2.13)	-0.08 (-0.23)
Far-Western	0.75 (0.47)	0.27 (0.12)	0.19 (0.03)	0.07 (0.01)	3.49 (1.90)	-0.06 (-0.33)
Per capita expenditure:						
1 (poorest quartile)	0.65 (0.40)	0.23 (0.06)	0.11 (0.03)	0.06 (0.00)	2.63 (1.38)	-0.33 (-0.49)
2	0.77 (0.56)	0.30 (0.15)	0.11 (0.06)	0.05 (0.01)	3.51 (2.18)	-0.01 (-0.18)
3	0.78 (0.73)	0.29 (0.29)	0.23 (0.09)	0.08 (0.02)	3.64 (3.09)	0.04 (0.18)
4 (wealthiest quartile)	0.86 (0.78)	0.51 (0.44)	0.28 (0.32)	0.24 (0.11)	4.69 (4.65)	0.45 (0.68)
War intensity:						
1 (lowest quartile)	0.71 (0.66)	0.42 (0.33)	0.26 (0.25)	0.19 (0.10)	3.97 (3.49)	0.12 (0.27)
2	0.73 (0.58)	0.31 (0.23)	0.19 (0.12)	0.11 (0.03)	3.47 (2.69)	-0.01 (0.04)
3	0.79 (0.59)	0.32 (0.20)	0.13 (0.09)	0.06 (0.01)	3.47 (2.43)	0.02 (-0.09)
4 (highest quartile)	0.78 (0.55)	0.24 (0.17)	0.12 (0.04)	0.06 (0.00)	3.28 (2.21)	-0.11 (-0.19)

Means shown for each category in 2004 (1996 in parentheses). This sample comprises individuals aged 6-17 in the cross-sectional data of NLSS I and II. Net enrollment is the number of students enrolled in a level of schooling who belong in the relevant age group, as a percentage of the population in that age group. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. War intensity is the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991.

Table 2 - War Intensity and Schooling Attainment (Cross Section)

Dependent variable:	Years of schooling				Grade-for-age z-score			
	2004 cross section		Pooled cross sections		2004 cross section		Pooled cross sections	
	OLS (1)	DID (2)	OLS (3)	DID (4)	OLS (5)	DID (6)	OLS (7)	DID (8)
Casualties (per thousand)	-0.227***		-0.253***		-0.056***		-0.065***	
	[0.059]		[0.054]		[0.017]		[0.015]	
Affected cohort: Aged 6-17		5.815***		-0.104		0.053		-0.038
		[0.279]		[0.142]		[0.072]		[0.041]
Casualties (per thousand) × Affected cohort		0.278**		0.225**		0.060*		0.058*
		[0.108]		[0.097]		[0.033]		[0.033]
Individual and household controls, and cohort fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Year fixed effects	N	N	Y	Y	N	N	Y	Y
District fixed effects	N	Y	N	Y	N	Y	N	Y
Mean of dependent variable	3.717	3.717	3.274	3.274	0.000	0.000	0.000	0.000
Std. dev. of casualties (per thousand)	0.686	0.686	0.723	0.723	0.686	0.686	0.723	0.723
Effect on dependent variable per std. dev. change in casualties	-0.156	0.191	-0.183	0.163	-0.038	0.041	-0.047	0.042
Number of observations	17479	17479	33464	33464	17479	17479	33464	33464
R-squared	0.46	0.47	0.45	0.47	0.27	0.30	0.27	0.29

Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(2) and (5)-(6) use only the 2004 cross section of NLSS; columns (3)-(4) and (7)-(8) pools the 1996 and 2004 cross sections. Odd-numbered columns depict OLS estimates while even-numbered columns depict difference-in-differences estimates. By definition, individuals from the 1996 cross section belong to unaffected cohorts, regardless of age. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. Individual and household controls include gender, parental schooling attainment, per capita consumption, and caste.

Table 3 - War Intensity and Schooling Attainment (Panel)

Dependent variable:	Years of schooling				Grade-for-age z-score			
	Second wave only		Full panel		Second wave only		Full panel	
	OLS (1)	DID (2)	FE (3)	FE (4)	OLS (5)	DID (6)	FE (7)	FE (8)
Casualties (per thousand)	-0.260				-0.064			
	[0.179]				[0.053]			
Affected cohort: Aged 6-17		4.981***	3.754***	3.683***		-0.881	0.004	-0.007
		[1.005]	[0.303]	[0.306]		[0.932]	[0.087]	[0.089]
Casualties (per thousand) × Affected cohort		0.782**	0.061	0.068		0.233**	0.064	0.058
		[0.301]	[0.265]	[0.298]		[0.095]	[0.083]	[0.087]
Individual and household controls	Y	Y	Y	Y	Y	Y	Y	Y
Cohort fixed effects	Y	Y	N	N	Y	Y	N	N
Year fixed effects	N	N	Y	Y	N	N	Y	Y
District fixed effects	N	Y	Y	N	N	Y	Y	N
Time-varying district fixed effects	N	N	N	Y	N	N	N	Y
Individual fixed effects	N	N	Y	Y	N	N	Y	Y
Mean of dependent variable	4.086	4.086	3.309	3.309	0.006	0.006	0.017	0.017
Std. dev. of casualties (per thousand)	0.593	0.593	0.593	0.593	0.593	0.593	0.593	0.593
Effect on dependent variable per std. dev. change in casualties	-0.154	0.464	0.036	0.040	-0.038	0.138	0.038	0.034
Number of observations	2991	2991	5982	5982	2991	2991	5982	5982
R-squared	0.53	0.59	0.92	0.93	0.31	0.38	0.84	0.85

Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(2) and (5)-(6) use only the second wave of the panel, where columns (1) and (5) depict OLS estimates, and columns (2) and (6) depict difference-in-differences estimates. Column (3)-(4) and (7)-(8) use both waves of the panel, where columns (3) and (7) augment difference-in-differences with individual fixed effects, and columns (4) and (8) augment difference-in-differences with individual fixed effects and time-varying district-level fixed effects. By definition, observations from the first wave in 1996 belong to unaffected cohorts, regardless of age. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. Individual and household controls include gender, parental schooling attainment, per capita consumption, and caste. Time-varying household per capita consumption is controlled for in the augmented difference-in-differences regression.

Table 4 - Displacement Propensity

Dependent variable:	Displacement during war					
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)
Skills	-0.006	0.111**	0.040***	-0.076**	-0.205	-0.020
	[0.014]	[0.045]	[0.009]	[0.031]	[0.126]	[0.030]
Landholding	-0.067***	-0.067***	-0.056***	0.012	0.023	-0.004
	[0.014]	[0.014]	[0.014]	[0.022]	[0.023]	[0.024]
Casualties (per thousand) × Skills				0.263***	1.699***	0.254***
				[0.086]	[0.564]	[0.080]
Casualties (per thousand) × Landholding				-0.177***	-0.220***	-0.040
				[0.056]	[0.057]	[0.056]
Casualties (per thousand) × Landholding × Skills				-0.215**	-1.300**	-0.232***
				[0.086]	[0.569]	[0.078]
Individual and household controls, and cohort fixed effects	Y	Y	Y	Y	Y	Y
District fixed effects	Y	Y	Y	Y	Y	Y
Number of observations	17657	17657	17657	12644	12644	12644
R-squared	0.24	0.24	0.24	0.25	0.25	0.26

Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(6) depict OLS estimates using the full 2004 cross section. Displacement during war is a dummy that equals to one if an individual migrated in 1996-2003. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. The three measures of skills are (i) parental secondary schooling - a dummy for father's attainment of secondary schooling [columns (1) and (4)], (ii) skilled occupation dummy [columns (2) and (5)], and (iii) non-agricultural occupation dummy [columns (3) and (6)]. Landholding is a dummy for agricultural land ownership. Individual and household controls include cohort, gender, schooling attainment, per capita consumption, and caste.

Table 5 - Maoist Abductions and Schooling Attainment (Panel)

Dependent variable:	Years of schooling				Grade-for-age z-score			
	Second wave only		Full panel		Second wave only		Full panel	
	OLS (1)	DID (2)	FE (3)	FE (4)	OLS (5)	DID (6)	FE (7)	FE (8)
Maoist abductions (per thousand)	-0.002				0.000			
	[0.006]				[0.002]			
Affected cohort: Aged 6-17		5.225***	3.736***	3.662***		-0.811	0.015	0.001
		[0.937]	[0.240]	[0.245]		[0.909]	[0.067]	[0.069]
Maoist abductions (per thousand) × Affected cohort		0.037***	0.013	0.015		0.011***	0.007**	0.006*
		[0.013]	[0.011]	[0.013]		[0.004]	[0.003]	[0.003]
Individual and household controls	Y	Y	Y	Y	Y	Y	Y	Y
Cohort fixed effects	Y	Y	N	N	Y	Y	N	N
Year fixed effects	N	N	Y	Y	N	N	Y	Y
District fixed effects	N	Y	Y	N	N	Y	Y	N
Time-varying district fixed effects	N	N	N	Y	N	N	N	Y
Individual fixed effects	N	N	Y	Y	N	N	Y	Y
Mean of dependent variable	4.086	4.086	3.309	3.309	0.006	0.006	0.017	0.017
Number of observations	2991	2991	5982	5982	2991	2991	5982	5982
R-squared	0.53	0.59	0.92	0.93	0.31	0.38	0.84	0.85

Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(2) and (5)-(6) use only the second wave of the panel, where columns (1) and (5) depict OLS estimates, and columns (2) and (6) depict difference-in-differences estimates. Column (3)-(4) and (7)-(8) use both waves of the panel, where columns (3) and (7) augment difference-in-differences with individual fixed effects, and columns (4) and (8) augment difference-in-differences with individual fixed effects and time-varying district-level fixed effects. By definition, observations from the first wave in 1996 belong to unaffected cohorts, regardless of age. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Maoist abduction measures the number of abductions by Maoists and other non-state agencies in 2002-2006, normalised by the total district population in 1991. Individual and household controls include gender, parental schooling attainment, per capita consumption, and caste. Time-varying household per capita consumption is controlled for in the augmented difference-in-differences regression.

Table 6 - Cohort and Gender Analyses of War Intensity Effects (Panel)

Dependent variable:	Years of schooling			Grade-for-age z-score		
	Full	Male	Female	Full	Male	Female
	FE (1)	FE (2)	FE (3)	FE (4)	FE (5)	FE (6)
Affected cohorts:						
Casualties (per thousand) × Aged 6 in 2004	0.807	-0.586	1.485	0.026	-0.470	-0.043
	[1.477]	[3.581]	[1.746]	[0.445]	[0.892]	[0.668]
Casualties (per thousand) × Aged 7 in 2004	0.041	1.285	-1.335	0.212	0.415	-0.190
	[1.600]	[2.508]	[2.431]	[0.435]	[0.659]	[0.694]
Casualties (per thousand) × Aged 8 in 2004	-0.066	-1.410	2.607	-0.456	-0.671	-0.042
	[1.815]	[2.662]	[3.856]	[0.376]	[0.502]	[1.163]
Casualties (per thousand) × Aged 9 in 2004	0.199	-0.851	-2.377	-0.352	-0.497	-0.977
	[1.020]	[2.167]	[2.915]	[0.338]	[0.598]	[0.852]
Casualties (per thousand) × Aged 10 in 2004	-0.308	0.998	-0.209	-0.212	0.313	-0.393
	[3.356]	[3.297]	[3.037]	[0.718]	[0.828]	[0.996]
Casualties (per thousand) × Aged 11 in 2004	1.819	1.814	1.798	0.213	0.357	0.003
	[1.814]	[2.274]	[2.495]	[0.447]	[0.562]	[0.775]
Casualties (per thousand) × Aged 12 in 2004	0.415	0.158	0.208	0.128	0.136	-0.208
	[0.351]	[0.396]	[1.251]	[0.179]	[0.172]	[0.683]
Casualties (per thousand) × Aged 13 in 2004	1.081	-0.832	1.847	0.102	-0.476	0.085
	[1.309]	[3.222]	[1.509]	[0.324]	[0.689]	[0.472]
Casualties (per thousand) × Aged 14 in 2004	-0.074	0.768	-1.213	0.230	0.328	-0.144
	[1.506]	[2.170]	[2.417]	[0.404]	[0.500]	[0.672]
Casualties (per thousand) × Aged 15 in 2004	-0.257	-1.774	2.498	-0.486	-0.765	-0.032
	[1.714]	[2.578]	[3.786]	[0.369]	[0.467]	[1.125]
Casualties (per thousand) × Aged 16 in 2004	-0.520	-1.368	-3.074	-0.328	-0.459*	-1.021
	[0.728]	[1.043]	[2.752]	[0.199]	[0.260]	[0.787]
Casualties (per thousand) × Aged 17 in 2004	-0.182	1.835	-1.162	-0.018	0.534	-0.206
	[3.130]	[3.211]	[2.947]	[0.669]	[0.768]	[0.944]
Casualties (per thousand) × Aged 18 in 2004	1.401	1.526	1.922	0.425	0.434	0.552
	[1.131]	[1.394]	[1.901]	[0.327]	[0.400]	[0.595]
Casualties (per thousand) × Aged 19 in 2004	1.064	0.954	0.625	0.388	0.369	0.231
	[0.868]	[1.107]	[1.895]	[0.250]	[0.282]	[0.560]
Casualties (per thousand) × Aged 20 in 2004	-0.077	-1.736	0.025	0.045	-0.512	0.028
	[0.712]	[2.603]	[1.302]	[0.209]	[0.599]	[0.377]
Casualties (per thousand) × Aged 21 in 2004	-0.264	-0.063	-0.608	0.038	-0.018	-0.043
	[1.118]	[1.804]	[2.188]	[0.293]	[0.401]	[0.537]
Casualties (per thousand) × Aged 22 in 2004	-1.260	-2.059	-0.532	-0.233	-0.469	-0.152
	[1.120]	[2.221]	[1.019]	[0.271]	[0.445]	[0.323]
Casualties (per thousand) × Aged 23 in 2004	0.394	-0.220	-1.679	-0.069	-0.184	-0.500
	[0.573]	[0.853]	[1.169]	[0.154]	[0.228]	[0.380]
Casualties (per thousand) × Aged 24 in 2004	-1.025	1.071	-2.506	-0.216	0.317	-0.510
	[2.267]	[1.877]	[2.659]	[0.472]	[0.339]	[0.588]
Casualties (per thousand) × Aged 25 in 2004	0.039	-0.094	0.160	0.058	0.010	0.054
	[0.207]	[0.761]	[0.276]	[0.094]	[0.184]	[0.106]
Unaffected cohorts:						
Casualties (per thousand) × Aged 26 in 2004	1.370	1.938	0.614	0.122	0.271	-0.020

	(1)	(2)	(3)	(4)	(5)	(6)
	[1.333]	[1.390]	[2.056]	[0.259]	[0.253]	[0.440]
Casualties (per thousand) × Aged 27 in 2004	0.017	0.096	-0.217	0.015	0.082	0.052
	[0.420]	[0.620]	[0.775]	[0.113]	[0.132]	[0.213]
Casualties (per thousand) × Aged 28 in 2004	-0.799	-0.473	-0.614	-0.106	-0.020	-0.164
	[0.607]	[1.031]	[0.919]	[0.154]	[0.161]	[0.278]
Casualties (per thousand) × Aged 29 in 2004	-0.183	-0.165	-0.215	0.056	-0.097	-0.032
	[0.494]	[1.468]	[0.565]	[0.141]	[0.256]	[0.168]
Casualties (per thousand) × Aged 30 in 2004	0.219	-0.649	0.200	-0.036	-0.300	-0.001
	[0.243]	[0.786]	[0.297]	[0.102]	[0.208]	[0.159]
Individual and household controls	Y	Y	Y	Y	Y	Y
Year and time-varying district fixed effects	Y	Y	Y	Y	Y	Y
Individual fixed effects	Y	Y	Y	Y	Y	Y
Mean of dependent variable	3.189	4.339	2.033	3.309	4.463	2.141
Number of observations	5982	3009	2973	5982	3009	2973
R-squared	0.93	0.93	0.93	0.85	0.88	0.84

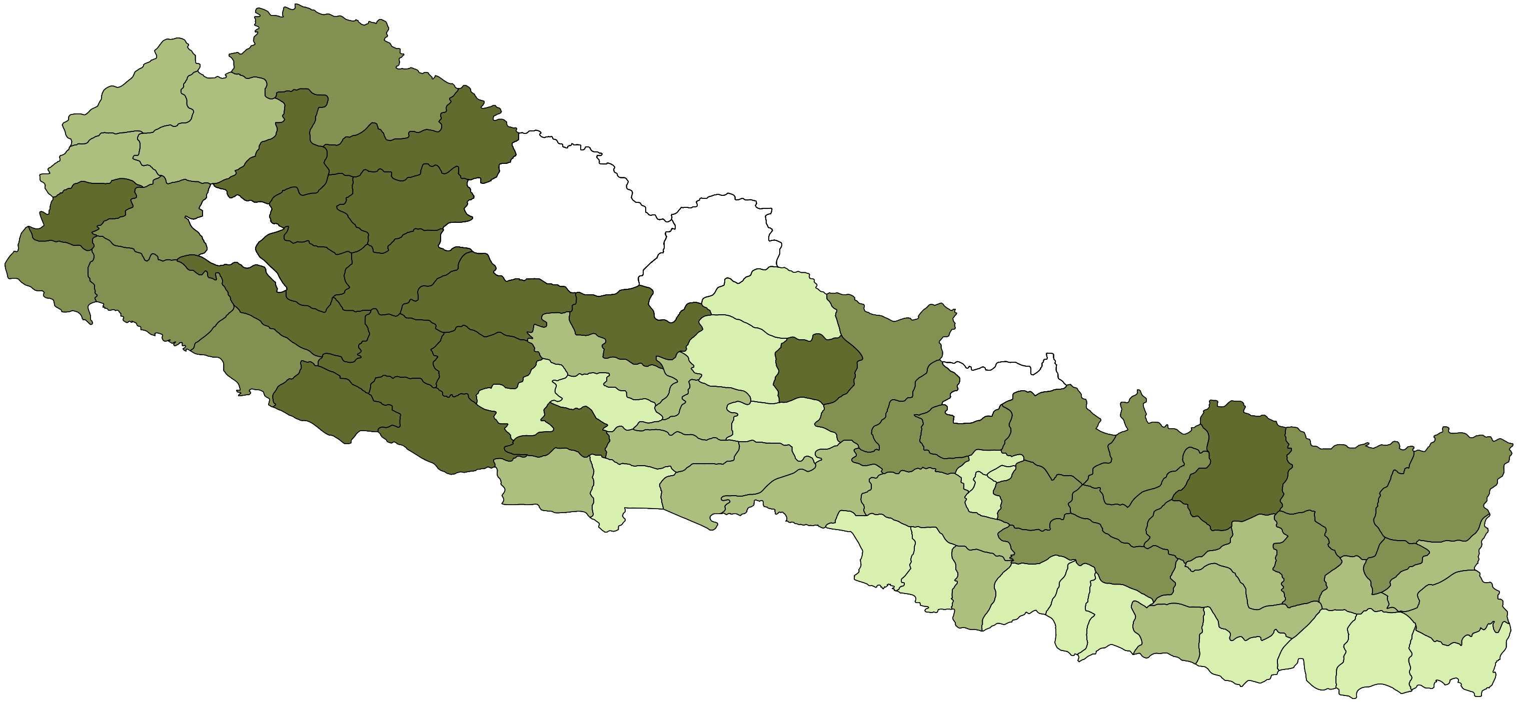
Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(6) use both waves of the panel and depict difference-in-differences estimates with individual fixed effects and time-varying district-level fixed effects. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. Individual and household controls include parental schooling attainment, per capita consumption, and caste.

Table 7 - Robustness Checks (Panel)

Dependent variable:	Grade-for-age z-score				Missing covariate(s)	Unbalanced panel	Achham district	Dolpa district
	FE (1)	FE (2)	FE (3)	FE (4)	FE (5)	FE (6)	DID (7)	DID (8)
Affected cohort: Aged 6-17	-0.007			-0.062	-0.304***	0.000	0.011	-0.003
	[0.089]			[0.093]	[0.034]	[0.000]	[0.007]	[0.004]
Affected cohort: Aged 6-18		0.006						
		[0.085]						
Affected cohort: Aged 6-19			0.012					
			[0.083]					
Casualties (per thousand) × Affected cohort	0.058	0.044	0.040	0.096	-0.047	0.000	-0.019*	0.002
	[0.087]	[0.087]	[0.085]	[0.088]	[0.034]	[0.000]	[0.011]	[0.003]
Individual and household controls	Y	Y	Y	Y	Y	Y	Y	Y
Cohort fixed effects	N	N	N	N	N	N	Y	Y
Year and time-varying district fixed effects	Y	Y	Y	Y	Y	Y	N	N
Individual fixed effects	Y	Y	Y	Y	Y	Y	N	N
Mean of dependent variable	0.017	0.017	0.017	-0.068	0.107	0.337	0.008	0.001
Std. dev. of casualties (per thousand)	0.593	0.593	0.593	0.603	0.593	0.593	0.723	0.723
Effect on dep. var. per std. dev. change in casualties	0.034	0.026	0.024	0.058	-0.028	0.000	-0.014	0.001
Number of observations	5982	5982	5982	5382	9142	9142	33464	33464
R-squared	0.85	0.85	0.85	0.85	0.72	0.99	0.04	0.02

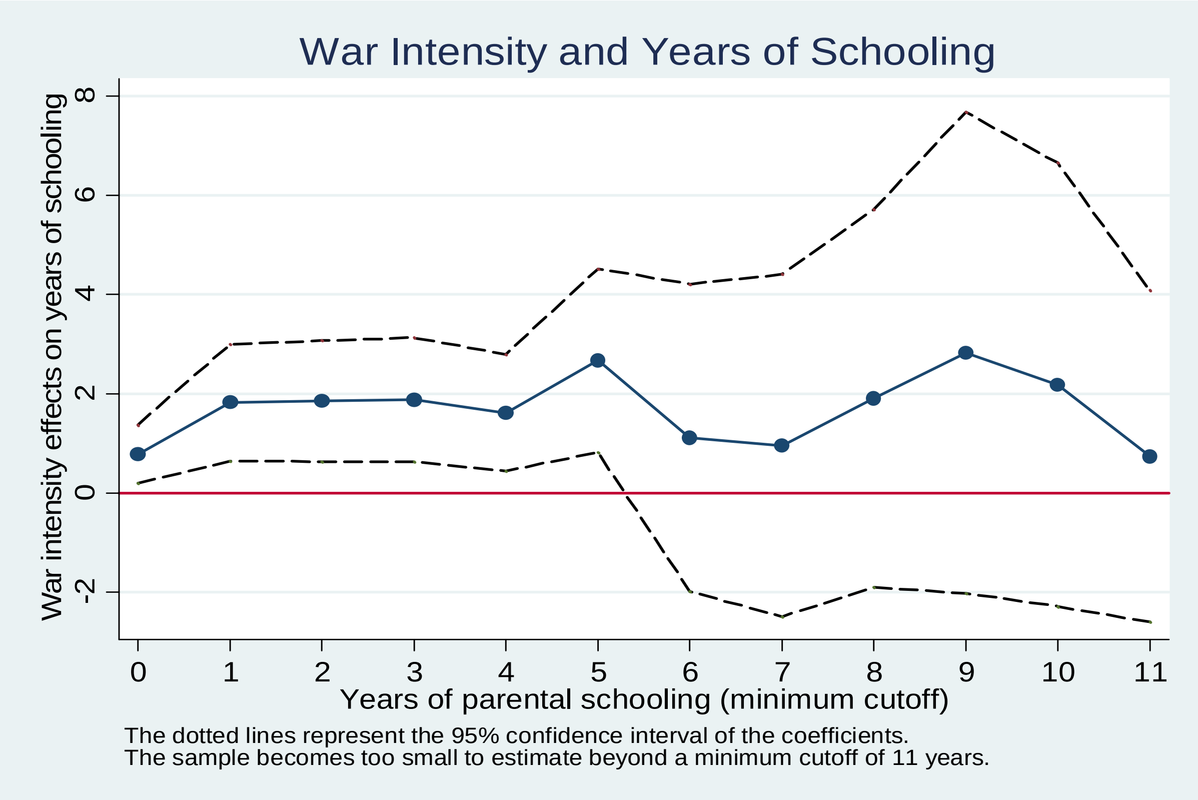
Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(6) use both waves of the panel and depict difference-in-differences estimates with individual fixed effects and time-varying district-level fixed effects. Columns (7) and (8) use the 1996 and 2004 cross section respectively, and depict difference-in-differences estimates. Dependent variables: missing covariate(s) equals one if an individual is missing at least one individual or household covariate; unbalanced panel equals one if an individual is observed only in one wave of the panel; Achham and Dolpa district dummies are self-explanatory. Observations in the district of Kathmandu are excluded in column (4). Not all individual and household controls are accounted for in column (5). By definition, observations from the first wave in 1996 belong to unaffected cohorts, regardless of age. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. Individual and household controls include gender, parental schooling attainment, per capita consumption, and caste. Time-varying household per capita consumption is controlled for in the augmented difference-in-differences regressions.

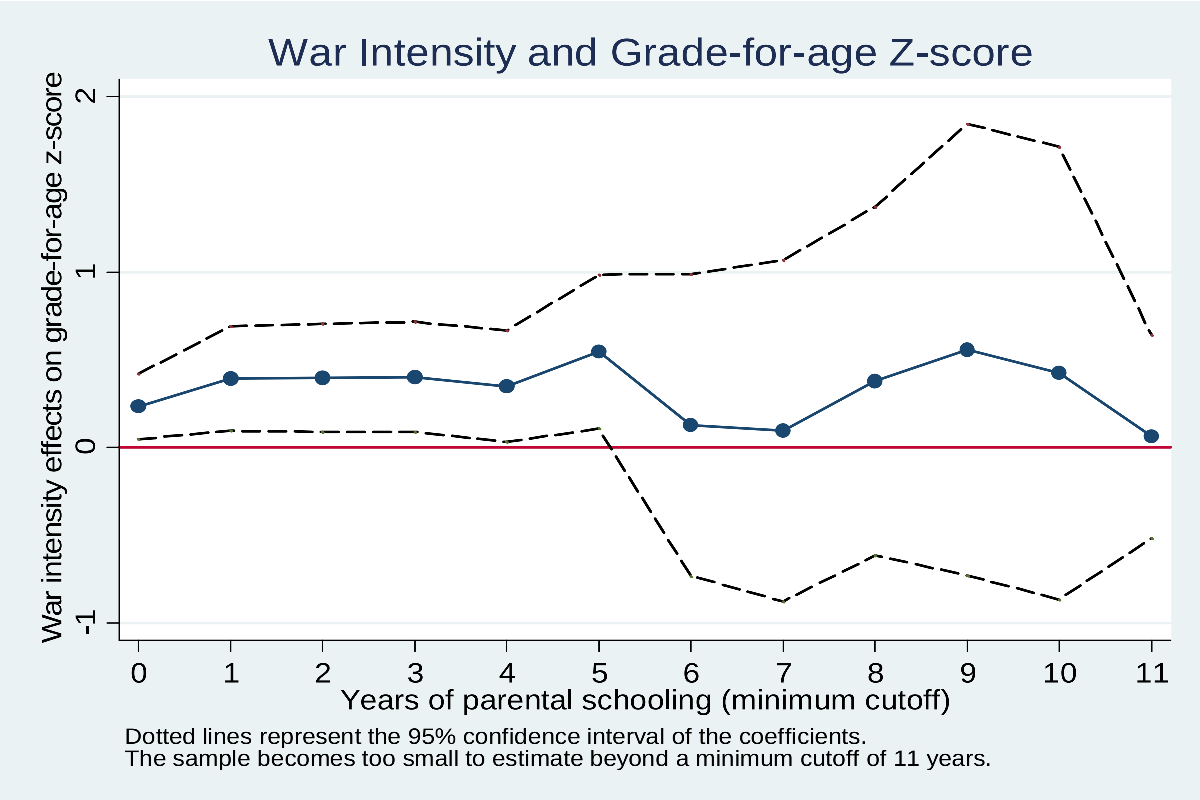
Figure 1 - War Intensity (By District)



Note: Shaded districts (71 of 75) are included the panel data, with darker shades denoting higher war intensity (in quartiles). The quartiles are, in number of casualties per thousand: 0-0.368, 0.369-0.650, 0.651-1.117, 1.118 and above.

Figure 2 - War Intensity Effects and Selection by Ability





Appendix Table 1 - Endogeneity of War Intensity

Dependent variable:	Casualties (per thousand)	Grade-for-age z-score	Casualties (per thousand)	Grade-for-age z-score	Casualties (per thousand)	Grade-for-age z-score	Casualties (per thousand)	Grade-for-age z-score
	OLS (1)	OLS (2)	SUR (3)	SUR (4)	OLS (5)	OLS (6)	SUR (7)	SUR (8)
Rural	0.064	-0.608***	0.064***	-0.608***	-0.101*	0.585***	-0.102***	0.585***
	[0.055]	[0.093]	[0.012]	[0.022]	[0.060]	[0.072]	[0.009]	[0.017]
Poverty rate	-0.600	-0.396**	-0.600***	-0.396***	-0.243	-0.258	-0.242***	-0.258***
	[0.559]	[0.164]	[0.029]	[0.053]	[0.481]	[0.198]	[0.030]	[0.058]
Literacy rate	-2.642**	1.248***	-2.640***	1.248***	-0.248	0.619**	-0.251***	0.619***
	[1.302]	[0.263]	[0.071]	[0.109]	[0.448]	[0.254]	[0.041]	[0.077]
Share of high caste	1.100*	0.014	1.098***	0.014	0.355	0.421***	0.354***	0.421***
	[0.596]	[0.130]	[0.035]	[0.053]	[0.435]	[0.133]	[0.030]	[0.048]
Maximum elevation	0.050**		0.050***		0.067***		0.067***	
	[0.020]		[0.002]		[0.022]		[0.002]	
Normalised road area	-0.067		-0.067***		-0.453**		-0.449***	
	[0.224]		[0.025]		[0.206]		[0.020]	
Region fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Breusch-Pagan test stat.			8.734	8.734			41.603	41.603
[p-value]			0.003	0.003			0.000	0.000
Number of observations	15985	15985	15985	15985	17433	17433	17433	17433
R-squared	0.59	0.17	0.58	0.17	0.54	0.16	0.54	0.16

Standard errors are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Results in columns (1), (2), (5), and (6) are separate OLS regressions; results in columns (3)-(4) and (7)-(8) are pairs of seemingly unrelated regressions (SUR). The left and right panels use the 1996 and 2004 cross sections respectively. The BreusGrade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. Covariates are defined as follows: rural is a dummy for rural district; poverty and literacy rates are district-level measures; share of high caste denotes the proportion of high caste individuals in the district; maximum elevation (in thousands of metres) is a terrain measure; normalised road area is the area used for roads relative to total district land area.

Appendix Table 2 - Displacement Propensity by Insurgent Influence

Dependent variable:	Displacement during war			
	Maoist districts only		UPF won seat(s) only	
	OLS (1)	OLS (2)	OLS (3)	OLS (4)
Parental secondary schooling	-0.019	-0.012	-0.005	-0.072*
	[0.020]	[0.046]	[0.021]	[0.041]
Landholding	-0.043	0.025	0.002	-0.006
	[0.028]	[0.044]	[0.025]	[0.028]
Casualties (per thousand) × Parental secondary schooling		0.015		0.056
		[0.063]		[0.152]
Casualties (per thousand) × Landholding		-0.087		0.001
		[0.069]		[0.100]
Casualties (per thousand) × Landholding × Parental secondary schooling		0.027		-0.003
		[0.069]		[0.151]
Individual and household controls, and cohort fixed effects	Y	Y	Y	Y
District fixed effects	Y	Y	Y	Y
Number of observations	5815	5815	2079	2079
R-squared	0.47	0.48	0.74	0.74

Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above in the 2004 cross section. Columns (1)-(2) depict OLS estimates for districts classified by the Nepalese government as being controlled by the Maoists; columns (3)-(4) depict OLS estimates for districts in which the United People's Front of Nepal (UPF) - the militant arm of the Communist Party of Nepal - won at least one constituent seat in the parliamentary elections of 1991. Displacement during war is a dummy that equals to one if an individual migrated in 1996-2003. Casualties (per thousand) is our measure of war intensity, calculated by the number of conflict-related deaths in 1996-2006, normalised by the total district population in 1991. Parental secondary schooling is a dummy for father's attainment of secondary schooling. Landholding is a dummy for agricultural land ownership. Individual and household controls include cohort, gender, schooling attainment, per capita consumption, and caste.

Appendix Table 3 - Falsification Tests

Dependent variable:	Years of schooling				Grade-for-age z-score			
	Second wave only		Full panel		Second wave only		Full panel	
	OLS (1)	DID (2)	FE (3)	FE (4)	OLS (5)	DID (6)	FE (7)	FE (8)
Rural dummy	-1.842***				-0.587***			
	[0.420]				[0.150]			
Affected cohort: Aged 6-17		4.138***	4.831***	4.580***		-1.159	0.115	0.020
		[1.026]	[0.407]	[0.394]		[0.932]	[0.147]	[0.165]
Rural dummy × Affected cohort		1.578***	-1.291***	-1.057**		0.503***	-0.088	0.009
		[0.463]	[0.488]	[0.453]		[0.148]	[0.161]	[0.175]
Individual and household controls	Y	Y	Y	Y	Y	Y	Y	Y
Cohort fixed effects	Y	Y	N	N	Y	Y	N	N
Year fixed effects	N	N	Y	Y	N	N	Y	Y
District fixed effects	N	Y	Y	N	N	Y	Y	N
Time-varying district fixed effects	N	N	N	Y	N	N	N	Y
Individual fixed effects	N	N	Y	Y	N	N	Y	Y
Mean of dependent variable	4.086	4.086	3.309	3.309	0.006	0.006	0.017	0.017
Number of observations	2991	2991	5982	5982	2991	2991	5982	5982
R-squared	0.55	0.59	0.93	0.93	0.34	0.39	0.84	0.85

Standard errors are in parentheses, and are clustered by sampling unit. * significant at 10%; ** significant at 5%; *** significant at 1%. This sample comprises individuals aged 5 and above. Columns (1)-(2) and (5)-(6) use only the second wave of the panel, where columns (1) and (5) depict OLS estimates, and columns (2) and (6) depict difference-in-differences estimates. Column (3)-(4) and (7)-(8) use both waves of the panel, where columns (3) and (7) augment difference-in-differences with individual fixed effects, and columns (4) and (8) augment difference-in-differences with individual fixed effects and time-varying district-level fixed effects. By definition, observations from the first wave in 1996 belong to unaffected cohorts, regardless of age. Grade-for-age z-score is the normalised grade attainment, relative to the individual's grade-for-age. Rural dummy is a binary indicator for rural districts. Individual and household controls include gender, parental schooling attainment, per capita consumption, and caste. Time-varying household per capita consumption is controlled for in the augmented difference-in-differences regression.