UNDERSTANDING THE DECLINE OF INTERPERSONAL VIOLENCE IN THE ANCIENT MIDDLE EAST

Giacomo Benati
Joerg Baten
Arkadiusz Soltysiak
Title: Understanding the decline of interpersonal violence in the ancient middle east

Abstract: How did human societies succeed in reducing interpersonal violence, a precondition to achieve security and prosperity? Given that homicide records are only available for the more recent period, much of human history remains virtually outside our purview. To fill this gap, a literature intersecting economics, archaeology, and anthropology has devised reliable methods for studying traumas deliberately inflicted in human skeletal remains. In this paper we reconstruct the early history of conflict by exploiting a novel dataset on weapon-related wounds from skeletons excavated across the Middle East, spanning the whole pre-Classical period (ca. 8,000-400 BCE). By documenting when and how ancient Middle Eastern populations managed to reduce intersocietal violence and achieve remarkable levels of development, we broaden historical perspectives on the structural factors driving human conflict.

JEL Codes: O15, N15, N35, N45, N55, P48

Keywords: Historical conflict, Middle East, preindustrial economies

Authors:
- Giacomo Benati
  Universitat de Barcelona
- Joerg Baten
  University of Tuebingen
- Arkadiusz Soltysiak
  Warsaw University

Email: giacomo.benati@ub.edu  Email: joerg.baten@uni-tuebingen.de

Email: a.soltysiak@uw.edu.pl

Date: June 2022
Understanding the Decline of Interpersonal Violence in the Ancient Middle East

J. Baten¹, G. Benati²*, A. Soltysiak³

Abstract: How did human societies succeed in reducing interpersonal violence, a precondition to achieve security and prosperity? Given that homicide records are only available for the more recent period, much of human history remains virtually outside our purview. To fill this gap, a literature intersecting economics, archaeology, and anthropology has devised reliable methods for studying traumas deliberately inflicted in human skeletal remains. In this paper we reconstruct the early history of conflict by exploiting a novel dataset on weapon-related wounds from skeletons excavated across the Middle East, spanning the whole pre-Classical period (ca. 8,000-400 BCE). By documenting when and how ancient Middle Eastern populations managed to reduce intersocietal violence and achieve remarkable levels of development, we broaden historical perspectives on the structural factors driving human conflict.

Main Text:

The explosion of a new large-scale conflict in eastern Europe forces us to reconsider once again the “long peace” hypothesis, i.e. that our era is considerably more peaceful than any other previous century (1, 2, 3, 4). Indeed, the prevalent claim that interpersonal violence (assault, murder, slavery, torture, despotism, cruel punishment, violent feuds, etc.) has been declining since the Enlightenment has been variously challenged (5) but so far, no consensus has been reached. In fact, because homicide statistics are only available for the recent period (3, 4), and because conflict records are subject to reporting biases (5), our understanding of long-term violence trends is still limited, the further back we go in time. This situation has generated highly conflictual or oversimplified narratives about the history of violence. A point in fact is given by Y. Harari’s bestseller “Sapiens” (6). According to Harari, scholarship is split into two factions, one that thinks that violence in human history only picked up after the Agricultural Revolution, while the other sees forager societies as exceptionally violent. In Harari’s view, the problem lies in the scarcity of comparative anthropological evidence about foragers and in the ambiguity of the archaeological evidence on conflict in early farming societies. Furthermore, studies based on more systematic analysis of historical and anthropological data on violent deaths simply infer that premodern (nonstate) societies were largely more violent than modern (state) ones (7, 8).

As a matter of fact, recent progress in forensic investigation techniques applied to human skeletal remains (9) prompted the proliferation of standardized databases of excavated skeletal data (10, 11), thus contributing tremendously to studying variations in human violence across different spatial, social, and temporal scales. We investigate in this article the causal factors driving the changes in long-term violence trends, building on an interdisciplinary literature – bridging economics, history, archaeology, and anthropology (12, 13, 14). Understanding differences in levels of violence in the past is not only central in historical social sciences, but it is also relevant for addressing present-day vulnerabilities (15). Indeed, since research on past human societies –

---

¹ Faculty of Economics and Social Sciences, University of Tübingen; Melanchtonstrasse 30, 72074, Tübingen, Germany. ² Faculty of Economics and Business, Department of Economic History, Institutions and Policy and World Economy, University of Barcelona (UB), Diagonal 690, 08034, Barcelona, Spain. ³Faculty of Archaeology, University of Warsaw; Krakowskie Przedmieście 26/28, 00-927 Warsaw, Poland. *Corresponding author Email: giacomo.benati@ub.edu
i.e. characterized by economies not dissimilar from present developing countries – may allow detecting factors that are not present in modern Western ones, providing to policymakers more specific information for addressing the consequences of the increasing instability in the most vulnerable areas of the world, hard pressed by Global Warming (12, 14, 15).

Fig. 1. Distribution of Middle Eastern archaeological sites that provided bioarcheological evidence for the present study.

Fig. 2. Distribution of observations in the study region.

Scope of the study
In this paper we report the findings of a study based on the construction and analysis of a novel dataset on violence-related injuries in seven Middle Eastern countries (Turkey, Iraq, Iran, Syria, Lebanon, Israel, and Jordan) for which studies of no less than 3,539 well-provenanced skeletons are available between Late Prehistory and the Classical age – i.e. c. 8,300-400 BCE – (16, 17, 18), a period that witnessed epochal changes, such as the earliest transition from foraging to farming.
the Urban Revolution, the rise of pristine states, cross-regional trade networks, and climate-induced crises (19). Through a standardized methodology for inferring weapon-related trauma in sufficiently well-preserved skeletons (Supplementary Materials), we find that interpersonal violence steadily declined during the Early and Middle Bronze Ages (ca. 3,000-1,500 BCE), when early states achieved substantial capacities (20, 21), and then increased again between the Late Bronze and the Iron Age (1,500-400 BCE), when epochal crises caused major upheavals across the whole region (22, 23).

Our results emphasize that even societies characterized by simple economic activities and limited access social orders – i.e. no secure monopoly over violence and elite control over assets and resources – could devise effective solutions for basic development problems, such as that of limiting violence (24). These findings provide three major advances over the related literature. First, we offer an interdisciplinary methodology for reliably assessing violence levels across the full spectrum of the historical experience – provided that sufficient bioarcheological data are available – thus greatly enhancing replicability. Second, we contribute to debunking the idea that most historical societies were simply locked into a Hobbesian “state of nature” where violence was endemic and where predation and coercion were the only strategies to gain power and wealth (25). Third, we evaluate arguably exogenous factors – i.e. climate shocks and imported military innovations – as violence predictors, reducing the concern for the potential biases generated by endogeneity or unobserved variables. Combined, these outcomes contribute to broadening historical perspectives on the drivers of conflict revealing lessons that may be useful for policymaking.

**Trends of violence in the early Middle East**

We assigned each observation to one of the six time periods, namely the (late) Neolithic period (from ca. 8,300 to 4,500 BCE), the Chalcolithic (ca 4,500 to 3,300), the Early Bronze Age (ca 3,300 to 2,000), the Middle Bronze Age (2,000 to 1,550), the Late Bronze (1,550 to 1,200), and finally the Iron Age (1,200-400). Following the extant research on violence, it is possible to single out many factors as potential covariates, such as state capacity, demography, income, inequality, trade, human capital, military technologies, climate, diet, geography, etc. (12, 14, 26, 27). We focus here on some key factors, especially the exogenous ones that can be reliably measured in the archaeological record, and we evaluate whether our estimates are affected by them or not. We control for settlement size (urban vs. rural; 28), the most important exogenous military innovations – i.e. iron technologies and mounted warfare (29) –, climate shocks – i.e. droughts (cold/dry spells) (30), and population fluctuations. We also include region fixed effects in our regressions to control for unobserved heterogeneity (see Table 2).
The empirical results clearly show that there was a much higher level of violence in the Chalcolithic period and then significantly elevated levels between the Late Bronze and the Iron Age (see Figure 3). Notably, the data from the Early and Middle Bronze Ages show a steady decrease in violence in the region. The elevated levels during the Chalcolithic and Iron Age periods are not unexpected. Indeed, the Chalcolithic period witnessed large-scale conflict which has been connected to the emergence of the first centralized proto states (31, 32). Similarly, the transition between the Late Bronze and the Iron Age experienced major upheavals – a 300-years drought that triggered large-scale population displacements, struggle over resources and state dissolutions, especially in the Levantine regions (33) – followed by the rise of superpowers, such as the Assyrian empire, that expanded through military campaigns, forced taxation and mass-scale deportations (23, 34). However, a substantial and continuous decline in lethal aggressions was experienced between 3,000 and 1,500 BCE, i.e. during the Early and especially Middle Bronze Ages. Notably, despite state formation/decline episodes, droughts, and conflict (35) – especially across the transition between the 3rd and 2nd millennia, when a shift towards drier conditions increased pressure on resources and triggered significant demographic changes (28, 36) – the Bronze Age Mesopotamian societies managed to reduce internal violence. They partly redirected conflict towards external enemies, as indicated by lower levels of trauma in Mesopotamia and by the match between the concentration of traumas and the royal inscriptions documenting the warfare campaigns of Mesopotamian kings along the eastern borders (17, 37).

An influential literature on historical violence trajectories (1, 3, 26) posits that the combined effect of the development of state capacity and trading systems is possibly the main force behind the declining homicide rates from the Middle Ages. Three key historical facts from our sample seem to support this idea (see also Supplementary Materials: Appendix A). First, although organized as city-states since around 3,000 BCE (39), it is only after 2,400 BCE that Mesopotamian states consolidate into centralized kingdoms first, and then into empires capable of
controlling internal violence (38, 39). Second, written law codes and legal remedies for private rights – enforced by formal courts – expanded the most between 2,100 and 1,700 BCE (20). Third, empirical analysis of long-distance exchanges indicate that trade expansion peaked between the late Early Bronze and the early Middle Bronze ages (21, 40, 41). Notably, together with the post-2,000 BCE “trade revolution”, cuneiform writing expanded dramatically both geographically and socially, indicating rising levels of literacy (42). These changes were flanked by the increasing professionalization of Mesopotamian armies (43), which possibly contributed to confine interpersonal violence to specific social arenas. Thus, state consolidation and the expansion of legal and commercial systems may have exerted an effect of lessening the propensity for violence in the sample.

Moreover, we assessed whether the changes towards less interpersonal violence during the Early and Middle Bronze ages might have been rather driven by inequality, climate, population fluctuations, or military technologies – and not by the introduction of the rule of law and the synergetic effects of lower interpersonal violence with trade. Consistently with our main hypothesis, we found that these alternative explanations do not seem to bear much weight. First, inequality – a possible driver of violence (44) – steadily increased during the Bronze Age (45), and it is therefore at odds with our findings. Second, when we formally assess the impact of droughts, demography, and exogenous military innovations (see Table 2 below), we only find that low population pressure in the post-4.2ka BP event may have contributed to the decreasing rates of lethal aggression in the Middle Bronze period – whereas controlling for population density indicates that the Chalcolithic, Late Bronze and Iron Ages were characterized by similar violence levels. Finally, while the spread of iron weapons somewhat aided the upward violence trend in the latter part of the sample, the introduction of mounted warfare had effect at all. Overall, this test, although limited to the violence covariates that are most efficiently measurable in the archaeological and historical records, suggests that it is difficult to envision mechanisms different from the ones we take into consideration.

Conclusions

Anthropologists and social scientists typically assume that states are effective in reducing interpersonal violence (1). This, however, is far from being granted for modern states, and certainly even less for historical ones (24). Furthermore, the emergence of early states in the ancient world is usually seen as shepherded by political elites that established “armed peace” via violent repression and absolutism (1, 24). This is not only paradoxical, since it implies exchanging lower war casualties with higher impositions and lower living standards, but also difficult to prove given the lack of violence time-series for much of early human history. By exploiting novel bioarchaeological evidence to quantify levels of interpersonal violence in the deep past, and exogenous sources of variation, we document more accurately when and in what circumstances human societies became effective in controlling interpersonal violence. We find that violence rates in the pre-Classical Middle East increased sharply when proto-states emerged in the Chalcolithic period, then decreased sensibly between the Early and Middle Bronze ages – when state development and trade expansion produced a strong “pacifying” effect – only to backslide between the end of the Bronze and the Iron Ages. Our analysis supports the idea that violence declines when state capacity increases. We advance the related literature by controlling for the impact of truly exogenous factors. Ultimately, we document that ancient Middle Eastern societies effectively solved a major development problem – reducing intersocietal violence – once early states achieved
core capabilities, i.e. territorial control, centralization of political violence, full-fledged conflict resolution systems, and actively supported the expansion of trade networks. However, security gains, aided by the positive effect of exogenous shocks temporarily reducing population pressure, were short-lived and the region witnessed major disruptions across the transition between the Bronze and Iron ages, possibly because of the rise of more extractive powers, more widespread access to cheaper weaponry, and mounting pressure over diminishing resources. Finally, the growing availability of anthropological and archaeological proxies, analyzed through rigorous econometric techniques, will provide not only new pathways for interdisciplinary cooperation between social sciences and historical disciplines, but also historical lessons for informing debates on social issues that are still debated today.

Methods

Measuring interpersonal violence

To study variations in violence levels in our sample (see Figure 1) we measure the intensity of interpersonal violence as the share of individuals exhibiting weapon-related wounds or cranial traumas (Supplementary Materials); following Baten and Steckel (26). The use of this indicator is motivated by the fact that in historical populations, cranial traumas and weapon injuries were often the result of interpersonal violence, as demonstrated by the systematic analysis of such indicators for a skeletal sample from European regions stretching 2,000 years (26). Hence, we do not agree with Harari (6) who provided the impression that archaeological evidence cannot inform us about early violence trends. To construct our database, we collected evidence of ante-mortem cranial traumas, excluding post-mortem cases. We also considered peri-mortem trauma, if it could be clearly distinguished from post-mortem damage. We only include traumas that were located above the hat-brim line – the line where an imagined hat brim touches the head, i.e. the frontal or parietal bone – which is usually the part of the skull hit by violent blows, in contrast to accidents that usually affect the area below the hat brim line. We included only skeletons aged 15 and above to remove bias relating to violence against subadults, which depending on cultural attitudes may have been stigmatized or not. In addition, we collected data about postcranial marks typically resulting from weapon wounds, such as arrow heads. This makes our study compatible with the Global History of Health project studies (26). However, the share of weapon wounds was tiny, most evidence comes from cranial trauma. We excluded stress or accident-related fractures. Furthermore, we considered “parry fractures”, i.e. fractures on ulnar bones that one person gets when using the arms to defend from an assailant (46). Although cranial and weapon traumas are obvious indicators of violent behaviors and allow us to measure the intensity of interpersonal violence in our sample, it must be stated that such a strategy typically underestimates violent deaths since a variety of wounds, particularly those to the soft tissues (e.g. poison, septic wounds, ruptured organs, etc.), do not leave marks on bones. Fortunately, Baten and Steckel (26) demonstrated that this data agrees well with other proxies for interpersonal violence, such as historical records on homicide, regicide, and elite deaths in battle – the latter include violence against soft tissues of course. We calculated trauma shares, which correspond to the ratio between the number of cranial and postcranial weapon wounds and the total number of sufficiently well-preserved skeletons from a given site/period. Our sample consists of 3,539 total skeletons examined with an overall number of traumas attested on 339 individuals, i.e. a weighted average share of 9.6%. As to regional distribution, we detect a higher occurrence of traumas from the peri-Mesopotamian regions – 8.4% from Turkey, 18.6% Iran, and 10.8% from Levant – and a lower incidence for Mesopotamia, 6.3%
After dropping observations for regions and periods with less than 30 observations, our figure 3 is based on 3,496 individuals (Table 1). Turning to potential selectivity biases that may affect our trends, we controlled for the distribution of observations between urban and non-urban (burial sites, rural villages, pre-urban communities, etc.) environments. We created dummy variables for special contexts, such as sacrifice of human retainers in royal interments. These variables allow us to adjust for religious and royal samples in the following regression-based estimates. Moreover, we identified clearly defined battle sites or cases of military attacks. Notably, these special contexts account only for 34 individuals over 3,549 observations, indicating that organized violence or special rituals cannot bias our results in a substantial way. In case of violence occurrence, the division between urban and rural contexts is considered meaningful by a substantial strand of historical literature (26, 28). Overall, our sample is quite well-balanced with 1,704 cases coming from urban contexts and 1,845 from non-urban ones, with this category grouping rural communities, burial sites, and pre-urban communities.

Table 1. Number of cases for violence shares by region and period. Regions: Iran; Levant (West Syria, Lebanon, Israel, Jordan); Mesopotamia (Iraq, East Syria); Turkey; Middle East, average. Periods (all BCE): Neolithic, 8300-4500; Chalcolithic, 4500-3300; Early Bronze, 3300-2000; Middle Bronze, 2000-1550; Late Bronze, 1550-1200; Iron, 1200-590. We excluded the number of observations that were below 30 cases per period and region for figure 3. These were 12 observations for Mesopotamia (Chalcolithic), 2 observations for Iran (Bronze Age), 5 and 9 observations for the Late Bronze Age for the Levant and Turkey respectively.

<table>
<thead>
<tr>
<th>Region</th>
<th>Neolithic</th>
<th>Chalcolithic</th>
<th>Early Bronze</th>
<th>Middle Bronze</th>
<th>Late Bronze</th>
<th>Iron</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iran</td>
<td>127</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>209</td>
<td>336</td>
</tr>
<tr>
<td>Levant</td>
<td>37</td>
<td>75</td>
<td>297</td>
<td>33</td>
<td></td>
<td>73</td>
<td>515</td>
</tr>
<tr>
<td>Mesopot.</td>
<td>56</td>
<td>90</td>
<td>66</td>
<td>51</td>
<td>51</td>
<td>314</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>1078</td>
<td>367</td>
<td>786</td>
<td>100</td>
<td></td>
<td>2331</td>
<td></td>
</tr>
<tr>
<td>Middle East</td>
<td>1171</td>
<td>569</td>
<td>1173</td>
<td>199</td>
<td>51</td>
<td>3331</td>
<td>3496</td>
</tr>
</tbody>
</table>

Empirical Approach

In the following we formally assess whether the changes over time that we observed were statistically significant. Moreover, we include control variables to analyze whether the changes towards less interpersonal violence during the Early and Middle Bronze ages might have been rather driven by climatic shocks, population pressure, or military technologies instead of the combination between rule of law and trade expansion. Our main control variables are defined as follows: “urban versus rural” character uses a threshold of ≥5 hectares of estimated area size. “Climatic shocks” are defined as rapid climate change events consisting of cool and dry spells (47), reducing temperature and precipitations, and generating droughts that negatively affected agricultural production (48). “Iron technology” is defined as the share of excavated sites that yielded iron materials and especially iron weapons. As this military technology was used first in the Caucasus region and then “imported” into Mesopotamia (29), it might have influenced our

---

2 We define “Levant” as the rain-fed region encompassing modern western Syria, Israel, and Lebanon; whereas “Mesopotamia” corresponds to the irrigated regions of modern Iraq and eastern Syria, up to the area of Tell Ashara (see Figure 2). Ancient Anatolia corresponds to modern day Turkey as ancient Iran to the modern country.
territory of study in a relatively exogenous way. Similarly, we add a control variable for the emergence of horse-mounted warfare, as again, horse-based military technology arrived from Central Asia and the Caucasus region and was not endogenously initiated in the region that we study (29). Together, these two innovations triggered a “military revolution” since, on the one hand, iron ores were widespread and iron weapons became quickly available for large sectors of the population, and, on the other hand, cavalry became crucial for warfare and changed the power ratio between mounted warriors and others (49). Finally, “population size” is included in the equation. We use the estimates by Palmisano et al. (28). As the area of our regions did not change, the changing population size estimates account for changes in population density and in high-density situations, for population pressure. Population size tended to increase, except after major epidemic disease events (as well as after climate shocks and institutional changes, but these factors we consider using separate variables). As violence might also have influenced population size, this might be a “proxy control” in Angrist and Pischke’s (50) nomenclature (see Supplementary Material, Appendix B). However, we need to include this variable in one of our estimates to meet the potential criticism that population pressure would otherwise not be taken into consideration.

Main results

In Table 2 we provide the regression results. The reference category for the time periods is the Chalcolithic, one period with tremendous violence as seen in Figure 3. It is not astonishing that we observe modest violence during the Neolithic period. During this period population density was very low therefore potentially conflicting groups could easily avoid each other. More notable is that – relative to the reference categories of the Chalcolithic – we observe a decline in violence during the Early Bronze and even more pronounced in the Middle Bronze. We explained this above with the expansion of the rule of law and the synergetic effect with trade, both innovations that are documented for these periods for the first time in human history, hence representing “natural experiments”. In contrast, while for the Late-Bronze and Iron Age the coefficients are also negative (hence less interpersonal violence than in the Chalcolithic), the coefficients are substantially smaller. The Mid-Bronze experienced a decline in violence by 15.3 percent, whereas the Later Bronze and the Iron Age only saw a 1.5 and 6.6 percent lower level of violence compared to the Chalcolithic (only the latter significant). Adding the control variable of an urban character of a site, we observe that the coefficients for Early and Mid-Bronze are unchanged. Urban sites had higher interpersonal violence during this period. In Table 2, column 2, we add the variable for climate shocks. The coefficient of this variable is statistically insignificant. The coefficients of Early and Mid-Bronze Age even become slightly larger in absolute size and keep their statistical significance, if the climatic shock variable is introduced. In Table 2, columns 3 and 4 we add iron and horse-riding technology which was relevant for military practices and might have had – via external effects – an indirect effect on interpersonal violence. We observe that the “iron” variable increases interpersonal violence rates as well, while the horse technology variable is insignificant and small. Again, the Early and Middle Bronze coefficient was not substantially affected. The coefficient of population density is positive (though not significant). Interestingly, the Late Bronze and Iron Age coefficients become insignificant, this suggests a similar violence level during the Late Bronze, Iron Age and Chalcolithic, if we control for the population density effect. Finally, we add a dummy variable for periods that did not experience a large-scale epidemic for half a millennium. Including this variable, we see that the coefficient for the Mid-Bronze declines substantially, hence a substantial share of Mid-Bronze interpersonal violence decline might have been caused by the fact that there was less population pressure, due to post- 4.2 ka BP demographic
crisis (28). But still, the coefficient for the Early and Mid-Bronze is statistically significant and substantial, suggesting that this period was different from the Chalcolithic, most likely because of the expansion of legal systems and the reinforcing effects of trade growth.

Table 2. Regression results. Did other determinants of violence invalidate the violence decline during the Early and Middle Bronze Age?

Note: *, **, *** refer to the 10; 5; 1 percent significance level, respectively. Robust standard errors are clustered at the region- and period level. The constant always refers to the Chalcolithic period. In column 1, the constant refers also to rural places. In column 2 the constant refers to the period of no climate shocks; in column 3, to the periods and places without documented iron tools/weapons. In column 4, the constant refers to periods and place without documented horse-riding technology use. In column 5, the constant cannot be interpreted, as population density was
never zero. Periods: Neolithic; Early Bronze Age; Middle Bronze Age; Late Bronze Age; Iron Age. See Table S1 for the description of each variable.

References


Acknowledgments: Funding by the German Science Foundation (SFB 1070 ResourcenKulturen) is gratefully acknowledged. We thank Alessandro Canci, William Pestle, J. Nicholas Postgate, Edward M. Luby, and Joyce Nassar for having shared with us relevant published and unpublished materials, and G. Chignoli for valuable research assistance.

Funding: German Science Foundation SFB 1070 “ResourcenKulturen” Grant.

Author contributions: Conceptualization: JB, GB; Methodology: JB, GB, AS; Investigation: JB, GB, AS; Visualization: JB, GB; Funding acquisition: JB; Project administration: JB; Writing – original draft: GB; Writing – review & editing: JB, AS
Supplementary Materials for

Understanding the Decline of Interpersonal Violence in the Ancient Middle East

Authors: J. Baten¹, G. Benati²*, A. Soltysiak³

Correspondence to: giacomo.benati@ub.edu

This PDF file includes:

Materials and Methods
Appendix A
Tables S1 to S3

Materials and Methods

The Data

Bone traumas as indicators of interpersonal violence. Our main measure of violence is the share of individuals with sufficiently well-preserved skeletal remains that present evidence of weapon-related wounds or cranial trauma. We focus on building the main indicator by counting the number of cranial traumas and weapon wounds relative to the total number of skeletons (1). The three categories of skull trauma – sharp, blunt, projectile – were merged into one for our purpose. Post-mortem damages were excluded. To construct our sample, we expanded the data on antemortem cranial traumas collected by Soltysiak (2) for 25 archaeological sites located in the Greater Mesopotamian region and stretching from the Neolithic to the modern age. We complemented his evidence by expanding toward the territories of what is today Iran, Turkey, Israel, Jordan, Lebanon and coastal Syria. This allows to include a total of 3,539 individuals, through an in-depth reexamination of paleopathological studies reporting both cranial and postcranial injuries, healed and unhealed, – e.g. blunt and sharp force traumas, projectile traumas, parry fractures of the ulna, etc. – for the Middle Eastern study regions.³ Cranial trauma and weapon wounds are obvious indicators of some components of overall violence. As a caveat, we should mention that the level of violence cannot, however, be assessed comprehensively by this measure because not all weapon wounds leave marks on bones (1). For example, lethal wounds to soft tissues, like a stab wound in the stomach, may not result in cut marks on bone surfaces, whereas other forms of killing, for example, poisoning, would not cause bone trauma at all. Assessing the level of violence by the occurrence of cranial trauma or bone marks is further complicated by the existence of both defensive and offensive weapon technology. For victims killed with some weapons such as a wooden stick instead of others such as a sword, it might be difficult to distinguish intentional violence from accidental trauma, especially if postcranial parts of the

³ Note that a similar study has been conducted by Cohen et al. (3) on a skeletal sample from the southern Levant, although they differentiate only three periods: between Chalcolithic and Iron Age (i.e., the overlap with our study is aggregated), Hellenistic to Byzantine Period, and Early and Late Arabic period (640 CE until 1917 CE). Wahl and Zäuner (4) criticized the methodology and results obtained by Cohen et al. (3), especially their assessment of representativeness.
skeleton are affected. Baten and Steckel (1) account for the omitted methods of killing by comparing cranial trauma and weapon wounds with other indicators, namely homicide, regicide (the killing of kings) and nobilicide (the killing of noblemen), for which they found very similar trends and cross-sectional differences. These other proxies of human violence include killing with poison, stomach stabbing and so on.

The special case of cranial trauma has been identified as very often being a result of interpersonal violence or inter-societal conflicts (3, 5). When Krakowka (6) analysed skull trauma in comparison with the coroners’ reports for the medieval city of London, she found that a large share of all homicides is directly related to skull trauma.

Following Wahl and Zäuner (4), several aspects need to be taken into account when proxying violence with evidence of trauma: 1) The skeletal sample studied for the population who lived at that time should be examined to ensure representativeness. In this respect, the collected data should not be from special cemeteries (such as elite cemeteries) or be otherwise unrepresentative of the population that one seeks to accurately reflect. 2) Taphonomical survival biases of bones and trauma signs need to be considered since it may be the case that only a selection of bones, and therefore trauma evidence, survived until the excavation. Especially in soils in which bone may decompose quickly, the signals of trauma cannot be identified in poorly preserved remains. Trampling behaviour of animals and many other taphonomic agents that destroy bone also need to be considered because in this way they could have destroyed bones that were buried close to the surface. 3) Trauma needs to be differentiated into antemortem, perimortem or post-mortem (in this case damage), with the most important distinction being between the latter two because post-mortem trauma does not indicate interpersonal violence but damage to the skeleton after death. As post-mortem traumatic lesions can sometimes seem to be caused by perimortem violence, researchers – albeit with some measurement errors – identify post-mortem damage by the amount of collagen of the bone that was already lost at the time of the fracture (8). Post-mortem fractures, often different in colour from the undamaged surface of the bone, tend to propagate at right angles to the bone surface, but peri-mortal trauma occurs at an oblique angle (9). 4) Some causes of death, such as poisoning or soft tissue wounds, do not leave marks on the bones (4). 5) Finally, violent trauma must be differentiated from accidental trauma. For this latter issue, the ‘hat brim line rule’ is one example of a suitable strategy to assess cranial trauma (4), and similar strategies exist for other parts of the human body. Concerning Wahl and Zäuner’s (4) first remark, the skeletal sample used in the study covers the Middle Eastern regions extensively and representatively, and the underlying evidence has been studied using a consistent methodology (12). As to regional distribution, we detect the same pattern already noted by Soltysiak (2) with a

But while the evidence on general population homicide is only available from the thirteenth century onwards for a few cities in Europe, the evidence on regicide and nobilicide only covers elite groups of the population.

Baten and Steckel (1) assessed whether the potential bias arising from the osteological paradox formulated for health stress markers might also apply to trauma and rejected this. Similar results, including for earlier millennia: Steckel and Rose (8).

Another part of the body where marks on bones are likely to be caused by violence are the forearms, and in particular, so-called parry fractures (7). These fractures mainly occur on the left ulna because the majority of the population would attack using their right arm while victims would defend the attack from violating their head by raising their left arm (7). Parry fractures are often mentioned in bioarchaeological textbooks as a potential indicator of interpersonal violence, but they have not yet been quantitatively studied for large regions and time periods (1). Baten and Steckel did not include it because the total number of clearly identified parry fracture cases was very small, as they observed. One disadvantage of using parry fractures as a violence indicator is that falls may mimic true parry fractures (7).

However, to discriminate between falls and blows in modern settings, similar blunt force trauma needs to be combined with other criteria associated with soft tissue and post-craniocranial injury patterns (10, 11).
higher occurrence of traumas from the peri-Mesopotamian regions – 200 from Turkey, 56 Iran, and 56 from Levant (West Syria, Israel, Lebanon) – and a lower incidence for Mesopotamia (Iraq, East Syria, 19 traumas). This regional distribution correlates well with historical conflict patterns indicating, first, that core Mesopotamian powers effectively centralized violence early on, and second, that inter-polity conflict was mostly concentrated in the border areas between Mesopotamian powers and neighboring regions (13, 14, 15).

To minimize regional selectivity biases, we ensured that sites from all Middle Eastern regions across all main time periods were included and a diversity of topographical features (sites located at the coast and in the interior, for example), as well as a representative urban/rural composition, were secured. To account for social selectivity biases, we controlled for the distribution of observations between urban and non-urban (burial sites, rural villages, pre-urban communities, etc.) environments and we created dummy variables for special contexts, such as sacrifice of human retainers in royal interments and for battle sites or cases of military attacks. These special contexts account only for only 34 individuals over 3,549 observations, i.e. 0.9%, indicating that organized violence or special rituals do not bias our sample. In case of violence occurrence, the division between urban and rural contexts is considered meaningful by a substantial strand of historical literature (1, 23). Indeed, our sample is quite well-balanced with 1,704 observations coming from urban contexts and 1,845 from non-urban ones, with this category grouping rural communities, burial sites, and pre-urban communities.

On the Wahl and Zäuner (4) issue (2) of survival of bones bias, we only collected information from samples characterized by sufficient levels of bone preservation to assure that this issue does not generate bias. Following Soltysiak (2), in almost all case studies it was possible to estimate the number of preserved cranial vaults using drawings of graves, preservation charts, or number of crania, for which sex was assessed using cranial characteristics, any measurements of the cranial vault were taken and any non-metric traits (i.e., sutural bones and parietal foramen) were scored. However, the quality of this estimation cannot be directly assessed, and it is therefore possible that complete but fragmented crania may not have been counted. Soltysiak (2) calculated (antemortem) cranial traumas share for skeletons for which at least 75% of the cranial vault was preserved. We then expanded this dataset by collecting information on unhealed perimortem cranial and postcranial weapon wounds from sufficiently well-preserved skeletons. In other words, we included only evidence for which sufficient bone of the skull was available to calculate the rates.

Moreover, on (3) the issue of post-mortem violence, perimortem trauma is usually distinguished from post-mortem damage in the paleopathological analyses from which we have collected evidence. For the potential neglect of some causes of death that do not leave obvious marks on wounds (4), we can rely on the correlations of our indicator of trauma and weapon wounds with other indicators of violence, such as the intensity of warfare as derived from royal

8 As to ritual killing contexts, Arslantepe VIb provided such evidence in the context of the “royal grave” which has yielded three young individuals likely sacrificed (16). Most famously, the Royal Cemetery of Ur provides evidence of individuals – likely royal retainers – seemingly killed to be interred with their sovereigns (17), and, lately, at Başur Höyük a large-scale interment provided possible evidence of ritual killings connected to the burial of a high-ranking individual (18). Turning to battle sites, the most striking evidence of warfare episodes come from the skeletons excavated in the Halzi Gate at Nineveh (19), attesting the demise of the Assyrian empire in 612 BCE at the hands of the Medes and Babylonians armies, and the ones from the citadel of Hasanlu (20), victims of an attack possibly perpetrated by an Urartian army that sacked and destroyed the period IV site (9th century BCE; 21). A possible violent attack, although of a smaller scale, seem to be attested in the skeletons from the MBA village of Kaman-Kalehöyük (22).
inscriptions and historical accounts (results upon request). Finally, (5) on the distinction between violent and accident trauma, the study of Baten and Steckel (1) calculated the correlation between (a) the share of violent trauma above the hat brim line and (b) the share of all cranial trauma. Given that the correlation was strong for the large European sample, we would expect the same for our Middle Eastern sample. In sum, the potential biases of bone survival, post-mortem trauma, other killing methods and accident trauma are not likely to apply to our sample.

In conclusion, the often violence-related cranial traumas, as well as weapon wounds, are now an established research instrument to assess the differences in interpersonal violence between time periods, regions, and groups. Although other parts of the body may show marks as well, according to earlier studies, the likelihood that these were caused by violence is lower than that of the former. Lower limb trauma, for example, is one of the most frequent traumas of all but often the result of accidents rather than of violence (25, 26).

**Violence covariates.** We focus here on several key factors, especially the exogenous ones, that can be reliably measured in the archaeological record, and we evaluate whether our violence trends are affected by them or not. Technological change – albeit more irregular in preindustrial societies – is one of the fundamental drivers of long-term economic growth and has ramifying effects on conflict, state formation, and information systems (27). Ironically, iron weapons were already used during the “Bronze Age”, although it was not the most important metal yet. Iron and cavalry were particularly critical technologies that conferred crucial advantages triggering, one the one hand, widespread adoption, and on the other, sparked “arms races” among competitors, spurring new technological improvements (27). To illustrate, although metal weapons were introduced during the Chalcolithic period, copper and bronze technologies remained firmly in the hands of elites due to their high costs in terms of technological know-how – based on alloying different base metals – and raw material procurement patterns – copper and tin are both notably absent from the Mesopotamian cores (28, 29). On the contrary, iron-smelting, due to the much more widespread availability of iron and its relatively simpler technology had a major impact on the evolution of technologies, as this strong and malleable material served as an input for a host of important technologies, military and otherwise (27). Similarly, although animal transport was first introduced during the Chalcolithic period (30), it is only towards the mid Early Bronze that animal-driven chariots were widely used as elite weapons across the Middle East. The importance of the war-cart diminished across the transition between Early and Middle Bronze (31), until the introduction of the spoke wheel and the horse-driven light chariot eventually led to the empowerment of the cavalry, the so-called “Cavalry Revolution” (32). Although riding is attested on visual media from Mesopotamia since the late 3rd millennium BCE (33), the development of effective horse-riding in the Pontic-Caspian steppes, combined with the use of bows and iron weapons, triggered a military revolution that spread from the steppes south to the belt of farming societies throughout the first millennia BCE (27). Consequently, in contrast to the strongly centralized bronze weapons, the widespread accessibility of iron arms quickly made available to large sectors of the population effective and durable weapons, whereas cavalry became crucial for

---

9 Baten and Steckel (1) follow Kremer et al. (24) regarding the distribution of traumas below and above the hat brim line. The latter argued that cranial trauma located above the hat brim line – the line where an imagined hat brim touches the head – is probably caused by violence that more often affects this part of the skull (the parietal or the frontal bones). Accordingly, accidents, such as falling, are more likely to cause cranial trauma below the hat brim line (24). Baten and Steckel (1) found a close correlation between violence evidence based on the hat brim line rule and all cranial trauma (see the scattergram and the correlation coefficient in Appendix B).
war tactics and changed the power ratio empowering horse owners at the detriment of other social sectors (34).

To control for the presence of iron and horse-mounted warfare in our sample we have exploited studies mapping the spread of such innovations from the Pontic area to the Middle Eastern countries based on the one hand, on the retrieval of iron items/iron-smelting technologies in the archaeological record (35), and, on the other, on paleogenetic and faunal studies and pictorial representations on visual media for mapping the distribution of domesticated equids (33, 36, 37).

As other main exogenous determinants of violence, we consider two partly related factors: Climate shocks and population pressure. Turning to climate shocks, we control for the impact of Rapid Climatic Change events – i.e. droughts –, during the Holocene, i.e. after 11.700 BCE. Indeed, according to the paleoclimatic record, in our timeframe the Middle East was subjected to at least eight sharp climatic shifts – i.e. Younger Dryas, 10.2 k, 9.2 k, 8.2 k, 5.2 k, 4.2 k, 3.2 k, 2.7 k cal. BP events (38, 39, 40) that, unlike gradual climate changes, had major impacts on populations heavily reliant on their environments for survival. More specifically, adverse climatic conditions – i.e. extended periods compounding cold temperatures and reduced precipitation regimes (41) – have been variously linked to unrest in the ancient Middle East (42, 43) and in other areas of the premodern world (44). Indeed, the conflict literature finds that climate-driven adverse economic conditions affect the severity and intensity of conflict, especially in agriculturally dependent regions (44). The two potential pathways are via direct resource scarcity and physiological/psychological stress, or indirectly via reduced economic outputs, raising food prices and triggering unmanaged migrations (but weakening political institutions and rapid urbanization/deurbanization may be other explanations; 44, 45). Consequently, to assess RCC as potential determinant of violence trends in our sample we have collected information about the timing and duration of droughts during the Holocene from paleoclimatic studies and then we created a variable –RCC – that corresponds to the percentage of time over which, each observation in our sample, experienced a drought.

Finally, we consider population size (and, by implication, population pressure) as violence covariate. One hypothesis might be that high population pressure could increase violence (1). In contrast, a recent literature combining archaeology and economic geography concludes that agglomeration-driven growth in pre-industrial societies could lead to substantial improvements in standards of living, spanning from material well-being, to declining inequality and lower interpersonal violence (46, 47). Notably, lower population density might result from major pandemics and other major disease waves (and/or natural disasters). Indeed, health conditions matter for aggregate economic and violence outcomes (48). Poor health and prevalence of infectious diseases may directly reduce productivity, life expectancy and investments in human capital, hindering economic development in a society and eventually triggering unrest. Although a historical epidemiological literature establishes a close correlation between climate shocks, disease outbreaks, unrest, and conflict (49, 50), an economic historical literature targeting the long-term impact of epidemics in the premodern world concludes that severe mortality crises, in the long run, could reduce inequality, triggering positive distributional effects and widespread welfare gains (51, 52). To test these competing hypotheses about the impact of population size on violence, we used the summed probability distribution (hereafter SPD) of calibrated radiocarbon dates estimated by Palmisano et al. (39), normalizing their data to average 100. SPDs of radiocarbon dates capture the density of human activities in each moment/place and are a new tool that archaeologists have recently adopted to reconstruct population trends in addition to more traditional archaeo-demographic proxies such as overall settlement counts and settlement sizes.
The overall agreement between SPDs and settlement-based proxies indicates that SPDs are a good instrument to reconstruct past population dynamics in the study sample (39). Indeed, in our sample we observe low levels of population growth across the Neolithic and Chalcolithic periods, followed by a proper take-off during the Early Bronze age, then a sharp drop during the Middle Bronze, a dramatic, but short-lived, increase during the Late Bronze and again a decrease during the Iron Age.

Appendix A. Evolution of Political, Legal and Trading Systems in the Ancient Middle East – Stylized Facts

*Neolithic* – For much of the Neolithic period, the social structure of the Middle Eastern communities was based on kinship with little stratification and limited socio-political hierarchy (53). Lacking written records, we can rely on comparative ethnographic evidence to suppose that some informal dispute resolution mechanisms were developed by these pre-state communities (54). Regional-scale networks of interaction have been reconstructed for some commodities chipped stones, obsidian, and other semi-precious goods (55, 56).

*Chalcolithic* – With the shift from village-based to urban lifestyles, the first forms of centralized leadership started to develop and eventually took the shape of proto-states – possibly controlling small territories around the main urban center – possible already around 4000 BCE (57). Soon after the introduction of writing in the southern alluvium, around 3350 BCE, written land deeds start to be attested, indicating the emergence of formal legal systems, in these early times only available to urban elites (58, 59). The first extensive long-distance networks were organized between early urban communities across the alluvium and mostly consisted of exchanges of goods for elite consumption (60).

*Early Bronze Age* – At the onset of the Early Bronze age period, early proto-states consolidated into city-states, first, and then morphed into territorial kingdoms towards the mid-3rd millennium BCE, a process driven by the emergence of formal kingship (61, 62). Finally, the end of the EBA also witnessed the rise (and fall) of empires capable of solidifying control of the alluvium and large portions of the surrounding areas, administered via extensive bureaucracies (63). In addition to state capacity, the latter EBA also witnessed an unprecedented expansion of trade and legal systems. To elaborate, long-distance exchanges peaked between 2500 and 2300 (64, 65), whereas legal remedies for nonelite groups expanded after 2500 BCE (66) together with the progressive formalization of legal systems as attested by the first legal codes issued by Mesopotamian rulers and by an increasing importance of courts for settling disputes (59).

*Middle Bronze Age* – In the aftermath of the collapse of empires, a more fragmented landscape arose with city-states and territorial entities often engaged in struggles for supremacy (67, 68). Notably, after 2000 BCE, property rights for nonelites expanded with the proliferation of written contracts enforceable by courts (66) and the frequent issuing of law codes and edicts (69). Simultaneously, a “trade revolution” was triggered by technological changes connected to the democratization of metal tools, leading to the creation of two major interlocking exchange networks, managed by private traders, that facilitated trade across the region (70, 71).
Late Bronze Age – The LBA witnessed first the rise of powerful Levantine states and the clash with Egypt, then dissolution of palace-based system in the Levant (53), possibly connected to extended drought that triggered famine, conflict, mass migrations and economic downfall (72, 73). During the latter part of the LBA trade started to shift towards the eastern Mediterranean basin, as attested by the intense exchanges between Levant and Egypt portrayed in the Amarna letters (74, 75). Notably, after 1700 BCE, despite a contraction of documentation from the southern alluvium, the small number of law codes issued and a concentration of judicial power in the hands of royal courts, might suggest a reduction of legal remedies for nonelites (76, 77).

Iron Age – The aftermath of the LBA crisis witnessed the rise of universalistic empires (68, 78). Indeed, up to ca 600 BCE, the Assyrian empire expanded through military conquest from its northern Iraqi core to the Mediterranean, Central Anatolia, Egypt, and the Gulf (40). Then, the Achaemenian dynasty further expanded its territorial control, at one point stretching from the Danube to the Indus River (79). Although territorial consolidation and extensive overland road networks facilitated exchanges and communication within the Achaemenid empire (80), the rise of Phoenician trading centers and of Aegean powers shifted the focus of trade to the Mediterranean regions, with a substantial take-off in the second half of the 1st millennium BCE (81). Turning to legal systems, although rulers of the first millennium issued virtually no legal codes, private legal documents suggest that the extent of legal remedies available to nonelites was like that of the Middle Bronze Age, especially in the Assyrian and Babylonian empires (82, 83).

Appendix B: What are “Proxy controls”?
The concept of “bad controls” was explained Angrist and Pischke (84). If control variables are variables that are themselves potentially outcome variables or dependent variables, the coefficient on the main variable of interest might be biased. As an example, Angrist and Pischke explain the case of the decision about a college degree, and which impact it has on the wage of an individual. The assumption is that the decision about the college degree is randomly assigned to the persons who decide about it. The question is, should the researcher include occupation as an explanatory variable, as the right-hand side variable? The answer is no, because occupation is also determined by the decision about the college degree completion. In an empirical example, the decline of the college degree coefficient that Angrist and Pischke observe, when occupation is included cannot be sensibly interpreted.

Interestingly, Angrist and Pischke distinguish “bad controls” and “proxy controls”, which are variables between good and bad controls. For example, in the study mentioned above, ideally an IQ of the eighth grade should be included to control for the innate ability of the pupils. However, it might be that such an IQ is not available, as it is often the case in empirical research. Instead, only a variable on an IQ measure at the time of the first job interview is available. However, this IQ is both influenced by the individual’s innate ability, that the researcher wants to control for, and their college decision, which would lead to endogeneity. Again, the coefficient of college degree completion is not precisely the true coefficient of interest. However, including this proxy control might be "be an improvement over an estimation with no control at all" (Angrist and Pischke 2009). If the college degree completion coefficient is positive, both with the IQ proxy control included and not included, the researcher might expect to find a true coefficient of interest between the two.

In our paper, population size is such a proxy control between good and bad, because we can control for an underlying variable that we cannot measure, namely pandemic events, which
would be a “good” control. At the same time, it is a bad control because societies with more interpersonal violence might suffer from population decline. By comparing the sign of the coefficient between including this proxy control variable and not including it, we can more safely assure that the early and mid-Bronze coefficient is significant. Similarly ambivalent are some of the other control variables, hence we term them “proxy controls”.

References


Table S1.
Variable descriptions and sources.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trauma share</strong></td>
<td>Ratio between the number of skeletons showing weapon-related bone traumas and the total number of skeletons subjected to paleopathological analyses from the sampled site for each observed period</td>
<td>Authors’ coding (see Data S1)</td>
</tr>
<tr>
<td><strong>Iron</strong></td>
<td>Dummy taking positive value if in the sampled site, over the observed time-frame, iron-smelting technologies were present</td>
<td>Authors’ coding (see Data S1)</td>
</tr>
<tr>
<td><strong>Horse</strong></td>
<td>Dummy taking positive value if in the sampled site, over the observed time-frame, horses were bred</td>
<td>Authors’ coding (see Data S1)</td>
</tr>
<tr>
<td><strong>Urban</strong></td>
<td>Dummy taking positive value if the dimension of the sampled site, over the observed time-frame, was ≥ of 5 hectares</td>
<td>Authors’ coding (see Data S1)</td>
</tr>
<tr>
<td><strong>RCC</strong></td>
<td>Percentage of time affected by a RCC event – cooling/drying episode – for each observation</td>
<td>Authors’ coding (see Data S1)</td>
</tr>
<tr>
<td><strong>Population size</strong></td>
<td>We use the SPD of radiocarbon dates estimated by Palmisano et al., normalizing their data to average 100.</td>
<td>(85)</td>
</tr>
</tbody>
</table>