Univerzita Karlova

Fakulta sociálních věd

Institut politologických studií

Ondřej Rosendorf

Campaign to stop ‘killer robots’: prospects of a preemptive ban on autonomous weapons systems

Diplomová práce

Praha 2018

Autor práce: Mgr. Ondřej Rosendorf

Vedoucí práce: PhDr. Michal Smetana, Ph.D.

Rok obhajoby: 2018

Bibliografický záznam

Rosendorf, Ondřej. Campaign to stop ‘killer robots’: prospects of a preemptive ban on autonomous weapons systems. Praha, 2018. 99 s. Diplomová práce (Mgr.) Univerzita Karlova, Fakulta sociálních věd, Institut politologických studií. Katedra bezpečnostních studií. Vedoucí diplomové práce PhDr. Michal Smetana, Ph.D.

Abstrakt

Tato práce se zabývá tématem autonomních zbraňových systémů a jejich možného preventivního zákazu s ohledem na mezinárodní diskuzi, která probíhá na různých multilaterálních fórech, včetně Rady pro lidská práva, Prvního výboru Valného shromáždění či Úmluvy o některých konvenčních zbraních na půdě OSN. Cílem práce je poskytnou rozsáhlý empirický přehled o podstatě těchto diskuzí a předpovědět nejpravděpodobnější výsledek diskuzí za pomoci odhadu preferencí států s využitím metod obsahové analýzy a teorému středního voliče.

Z teoretického hlediska vychází práce z defenzivního realismu, poznatků z teorie kontroly zbrojení, obchodu se zbraněmi, a částečně také z institucionalism, ze kterého si práce vypůjčuje koncept legalizace. Po metodologické stránce spoléhá práce na kvantitativní metody, zejména obsahovou analýzu k získání dat, a teorém středového voliče k předpovědi pravděpodobného výsledku diskuzí. Práce využívá také metodu regresivní analýzy k výzkumu aktivity států v rámci výše zmíněných mezinárodních fór.

Závěr práce shledává, že nejpravděpodobnějším výsledkem diskuzí na téma autonomních zbraňových systémů je mírně závazná forma hybridní regulace, která odpovídá specifickým řešení, jako jsou rámcová konvence či moratorium. Dalším zjištěním práce je, že aktivitu států v diskuzích na výše zmíněních fórech můžeme vysvětlit na základě podnětů a kapacit států k vývoji robotických a autonomních technologií spojených s autonomními zbraňovými systémy.

Abstract

This thesis addresses the issue of autonomous weapons systems and their potential preventive prohibition with regard to current international discussions at multilateral forums such as the Human Rights Council, First Committee of the General Assembly, and Convention on Certain Conventional Weapons at UN. The aim of this thesis is to provide an extensive empirical account of the substance of those discussions and their most likely outcome, estimating state preferences with use of content analysis and the likely outcome with median voter prediction.

From a theoretical standpoint, the thesis draws from defensive realism and contributions of arms control, arms trade as well as institutionalist literature from which it draws the concept of legalization. From a methodological standpoint, the thesis relies on quantitative methods, in particular, content analysis for collection of data and median voter theorem for prediction of the likely outcome. In addition, the thesis uses the method of regression analysis to examine states' activity at the aforementioned fora.

In conclusion, the thesis finds that the most likely outcome of discussions on autonomous weapons systems is a moderate-obligation form of hybrid regulation, which includes solutions such as framework convention and moratorium. Further finding of the thesis is that states' activity at the mentioned fora can be explained with their incentives and capacities for development of robotic and autonomous technologies related to autonomous systems.

Klíčová slova:

Autonomní zbraňové systémy, AWS, vražední roboti, kontrola zbrojení, CCW, preference, moc, regulace, zákaz

Key words:

Autonomous weapons systems, AWS, killer robots, arms control, CCW, preferences, power, regulation, ban

Rozsah práce: Text práce (úvod – závěr) čítá 216 486 znaků včetně mezer, což činí zhruba 125 normostran.

Prohlášení

1. Prohlašuji, že jsem předkládanou práci zpracoval samostatně a použil jen uvedené prameny a literaturu.
2. Prohlašuji, že práce nebyla využita k získání jiného titulu.
3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne .... Ondřej Rosendorf ....

Poděkování

Na tomto místě bych rád poděkoval vedoucímu práce PhDr. Michalovi Smetanovi Ph.D. za cenné připomínky k diplomové práci a za pomoc při realizaci mých akademických ambicí nad rámec této práce. Současně bych rád poděkoval Dr. Michaelovi C. Horowitzovi, Dr. Rebecce Crootofové, Dr. Frankovi Sauerovi a Paulovi Sharremu za feedback k práci.

Table of Contents

1. Introduction

The advent of the Campaign to Stop Killer Robots (CSKR) constitutes one of the more important developments in the field of arms control in recent years. With the movement of civil society activists and NGOs now backed by more than twenty countries, including China, expectations begin to build up with regard to potential prohibition of “killer robots”, or, more precisely, “autonomous weapons systems” (AWS) (CSKR 2018a; 2018b). Public figures such as Elon Musk, Stephen Wozniak, or recently deceased Stephen Hawking have called for a prohibition of AWS in the 2015 open letter published by the Future of Life Institute (FLI), which says: “If any major military power pushes ahead with AI weapon development, a global arms race is virtually inevitable, and the endpoint of this technological trajectory is obvious: autonomous weapons will become the Kalashnikovs of tomorrow” (FLI 2015). There are various reasons to believe that development and, eventually, use of systems operating without any human input could have grave consequences, not least due to legal, military, ethical, and operational challenges (see UNIDIR 2017a; Scharre & Horowitz 2015). However, much of the enthusiasm is unwarranted. It is probably unreasonable to expect countries to unconditionally agree to prohibition of systems with immense military benefits (see Arkin 2017; Marchant et al. 2011).

Although the many concerns over AWS development have been taken up to international forums such as the Human Rights Council (HRC), First Committee of the United Nations General Assembly (UNGA), Convention on Certain Conventional Weapons at the United Nations Office of Geneva (CCW), at the backdrop, there has been a push for increasing autonomy in weapon systems in most technologically advanced states (see Boulanin & Verbruggen 2017; Roff 2016; Scharre & Horowitz 2015). Beyond the remotely controlled armed drones, lethal AWS such as loitering munitions with autonomy in target selection and engagement have already been in use by some countries (Boulanin & Verbruggen 2017: 50–54). States such as China and the US have also conducted tests of semi-autonomous swarms of drones with advanced collaborative behavior (see Altmann & Sauer 2017; Lachow 2017; Kania 2017). Investments into artificial intelligence and robotic technologies for civilian and military purposes are sky-rocketing and some countries have expressed their ambitions to lead in this prospective field (see Boulanin & Verbruggen 2017; ILW 2017; Kania 2017). Of course, all this raises a question, how realistic the proposal of a ban actually is? The secretive nature of military R&D, greed but also legitimate security concerns of security-seeking countries must inevitably lead any student of International Relations (IR) to question the likelihood of success of such endeavor.

Nevertheless, "[...] the autonomy of weapon systems is accelerating largely outside of public and academic attention in the discipline of International Relations [...]" (Bode and Huelss 2018: 1–2). This fairly new, quite technical but immensely fascinating and important subject has been barely touched by IR scholars. Today's academic discourse on AWS is plagued by overly simplistic perceptions that have been fittingly dubbed "roboutopia" and "robocalypse" by Paul Scharre (2016) (see Scharre 2016). The very few contributions associated with IR usually cover certain security or arms-control-related aspects such as AWS' impact on strategic stability (see Altmann & Sauer 2017; Rickli 2017; Scharre 2017), preventive measures (see Altmann & Sauer 2017; Altmann 2013; 2009), or regional security (see Leys 2018; Koh 2017). Independently of these contributions, a fairly large body of policy and legal literature on AWS has also emerged recently (see Crootof 2016a; Geiss 2015; Anderson & Waxmann 2013; Heyns 2013; Schmitt & Thurnher 2013; HRW 2012; Marchant et al. 2011), suggesting regulatory alternatives to a ban. However, rarely do the contributions provide an assessment of regulatory solutions that focuses on feasibility with regard to security demands and interests of the powerful states or states with high stakes in the discussions at the aforementioned forums.

In fact, there is quite likely no regulatory remedy for "killer robots" unless the nations with capacities and incentives in AWS-related development agree to it. Hence, the aim of this thesis is to provide an assessment of the proposed solutions to AWS regulation with regard to states' preferences derived from the substantive content of the multilateral discussions. In addition, the purpose of this thesis is to provide an extensive empirical account of the discussions at HRC, UNGA and CCW, and to find what drives the seemingly sudden international interest in the topic. In this regard, the thesis adopts a defensive realist theoretical framework, which also incorporates certain contributions of the arms control and arms trade literature, or institutionalist literature, without necessarily compromising on the very essential assumptions of defensive realism (see Glaser 1995; Waltz 1979; Jervis 1978). This somewhat eclectic approach constitutes an advantage in addressing an issue multi-disciplinary in nature. In accordance with defensive realism, the author presumes that the plethora of security challenges associated with AWS development give rise to the security dilemma—that is, a situation in which an increase in one's security leads to decrease in security of others—which can bring states to cooperation, and in particular, through arms control agreements (see Tang 2009; Glaser 1997; Jervis 1978; Herz 1950).[1] The following two questions guide the thesis.

Research question 1: Which of the proposed regulatory solutions—ranging from no regulation to total prohibition—is likely to garner the most support from states involved in discussion on AWS?

Research question 2: What explains the activity of states at various fora, including the HRC, UNGA, and CCW where the discussions on AWS take place?

Accordingly, in addressing the two research questions, the thesis focuses on two dependent variables and two sets of relevant hypotheses. First, to find which of the proposed regulatory solutions is likely to prevail, the thesis looks at states' preferences, the first dependent variable. The current state of international debate and the very fact that it takes place prevailingly at forums for arms control seems to suggest that countries would rather have some regulation than no regulation at all. However, due to concerns over relative gains from distributional consequences of the agreement (see Fearon 1998; Glaser 1995; Grieco 1988), different capacities and incentives in AWS development as a result of the external environment as well as domestic processes responding to those external conditions (see Sterling-Folker 1997; Jervis 1999), and, particularly, difference in technological advancement (see Hansel et al. 2018), countries' preferences over specific regulatory solution to AWS different as well. Importantly, what form of regulation ultimately prevails depends on the process of bargaining and the power—in particular, material power and the

[1] This allows the author to draw upon the rich repertoire of proposed regulatory solutions to AWS found in the policy and legal literature on AWS, designing a tailor-made framework for assessment of states' preferences and drawing a connection between the two.

role of the context (Waltz 1979: 131; Schelling 1956: 282)—associated with states' preferences. The thesis suggests that one way to estimate preferences is to look at the substantive content of countries' statements at the various forums with respect to perceived challenges posed by AWS. In order to estimate states' preferences and find the most likely outcome, the thesis relies on methods of content analysis (see Krippendorff 2013; Neuendorf 2002) and a modified median voter theorem (MVT) model (see Bueno de Mesquita 2013) to predict the likely outcome using power in association with state preferences.

Second, to find what explains the activity of states at various forums such as the HRC, UNGA and CCW, the thesis explores countries' attendance of the meetings at the fora. Furthermore, the activity here serves as an approximation of countries' interest in the topic of AWS. The thesis suggests that the activity can be explained by the underlying capacities and incentives linked to AWS-related development. In other words, it would be ill-advised to assume that states' discussion on AWS and association to the CSKR cause are somehow altruistic. The thesis argues that most technologically advanced states have the greatest stake in discussions because any prospective agreement may affect them disproportionately (Hansel et al. 2018: 6). Countries therefore have a motivation to act through international institutions but also to signal benign intentions to ameliorate concerns over potential arms-racing (Glaser 2003: 409). The author suggests multiple variables reflecting the capacities and incentives, including external threats, research and development spending and military expenditure, existence of programs on armed drones and others that may affect countries' activity (Altmann & Sauer 2017: 118; Pamp & Thurner 2017: 460–461; Glaser 2000: 251; Buzan & Herring 1998: 83; Pearson 1989: 154–155). Thus, to test whether these factors have an effect, the author employs a linear regression model (see Monogan 2015; Allison 1999).

The thesis is structured as following. First, the thesis provides an overview of the empirical matter through the background on autonomy in weapon systems. This is a necessary first step due to the fact that the topic at hand is relatively new, and because it constitutes an important precondition for accurate content analysis. In other words, it would be impossible to assess countries' statements if we lacked the necessary background on what they actually talk about. Second, an overview of empirical literature follows. This chapter summarizes and categorizes the various contributions and approaches to the subject, focusing on specific solutions to AWS regulation, and providing a basis for assessing the array of states' preferences on a continuum. Third, theoretical framework is laid out, drawing from assumptions of defensive realism, including elaboration on relevant concepts such as the security dilemma, arms-racing or arms control, deriving hypotheses for testing with regard to two research questions. Fourth, the thesis provides rationale for case selection, including 12 most technologically advanced and powerful participants of debates on AWS and describes methods of analysis—content analysis, MVT prediction and regression analysis—as well as the data for analysis. Fifth, conclusions are drawn, with findings that a moderate "hybrid regulation" regulation is likely to prevail and that institutional activity is driven primarily by incentives or capacities connected with AWS-related development, and, in particular, the program of armed drones.[2]

Lastly, certain clarifications need to be made with regard to the initial thesis project. In addition to the aims and purposes already mentioned here, the diploma thesis project intended to explore the role of stigmatization in the debates on AWS, with a hypothesis that states utilizing stigmatizing strategies in the discussions are likely to favor prohibitive forms of AWS regulation. Instead of focusing squarely on the role of stigmatization, the thesis here incorporates the role of stigmatization in a tailor-made framework of preventive arms control criteria, designed to estimate states' preferences towards AWS regulation, which accounts for the role of ethical aspects of AWS regulation. Furthermore, the author also reformulated a hypothesis that states with limited AWS-related capabilities are more likely to support more restrictive forms of regulation. The thesis here chooses broader formulation which focuses on power rather than AWS-specific capabilities which would be difficult to estimate due to the lack of available data. Finally, the project also initially intended to explore the role of issue-linkage by employing a two-dimensional spatial model. Nevertheless, attempting to employ the role of issue-linkage would likely dilute the theoretical framework, and the lack of data for the potential second dimension too proved to be problematic.

[2] In other words, the institutional activity is highly positively correlated with existence of a program to develop armed drones also known as "unmanned aerial combat vehicles".

2. Background on Autonomy in Weapon Systems

The purpose of this chapter is to provide background information on AWS. Accordingly, the chapter is divided in three parts concerned with purpose and uses of autonomy, nomenclature, and classification of today's systems, as well as drivers and perils of the increasing autonomy, including various ethical, legal, security, and operational challenges of AWS (see Altmann & Sauer 2017; Arkin 2017; Boulanin & Verbruggen 2017; UNIDIR 2017a; Scharre & Horowitz 2015; Krishnan 2009). The reason for providing background information on the technology as such is twofold. First, the topic is relatively new, at least in the discipline of IR, and thus requires an introduction to the most fundamental issues that make it an important subject for research (Bode & Huelss 2018: 1–2). Second, the background information is also necessary for estimation of states' preferences of AWS regulation, specifically with regard to the method of content analysis. In order to translate country statements in policy preferences, it is essential to understand the substance of what they seek to communicate, first.

2.1. Substance of Autonomy: Definitional Approaches and State of the Art

First and foremost, there is no internationally accepted definition of AWS (Geiss 2016: 6; Crootof 2015: 1837; Scharre & Horowitz 2015: 3). Definitional contentions typically stem from the different perceptions of "autonomy". From a technological perspective, autonomy implies simply a capability of unsupervised operation, whereas philosophical perspective implies moral agency and free will (Arkin 2017: 35; Krishnan 2009: 43). However, most authors refer to autonomy from the fist point of view. For illustration, Scharre and Horowitz (2015) describe autonomy as "[...] the ability of a machine to perform a task without human input" (Scharre & Horowitz 2015: 5). Meanwhile, philosopher Robert Sparrow suggests that autonomy goes hand in hand with moral responsibility. According to his view, "[t]o say of an agent that they are autonomous is to say that their actions originate in them and reflect their ends" (Sparrow 2007: 65).

Sometimes, the term "autonomous system" is used to refer to AWS or non-armed version thereof. For example, UNIDIR (2016) defines autonomous systems as systems that "[...] operate without human intervention in the physical world or some kind of virtual environment" (UNIDIR 2016: 4). Thus, the term system refers to the means through which autonomy manifests. Similarly, Scharre and Horowitz define autonomous system as "[...] a machine, whether hardware or software, that, once activated, performs some task or function on its own" (Scharre & Horowitz 2015: 5). Typically, the term is used to refer to robots and unmanned vehicles but various machines used in everyday life also have certain automated functions, and could be described as autonomous systems. These include anything from Roomba vacuum cleaners to collision avoidance systems in military aircraft (Scharre 2015: 8–9; Scharre & Horowitz 2015: 5).

Terms such as "robot", "unmanned vehicle" (UV), and "artificial intelligence" (AI) are often used when referring to AWS. Krishnan (2009) describes robot as a "[...] machine, which is able to sense its environment, which is programmed and which is able to manipulate or interact with its environment" (Krishnan 2009: 9).[3] UV can be defined as "[...] a powered vehicle that does not carry a human operator, can be operated autonomously or remotely, can be expendable or recoverable, and can carry a lethal or nonlethal payload" (DoD 2007: 1).[4] While the term robot and UV typically refer to hardware, AI refers to software. Krishnan (2009) described AI as "[...] software that equips computerized system [with] human-like capabilities such as pattern recognition, text parsing and planning/problem-solving" (Krishnan 2009: 3).[5] In general, almost every AWS with a physical form can be described as a robot or UV with some a very "narrow" AI.

Autonomy in weapon systems is usually described as a "spectrum" or scale rather than a binary (UNIDIR 2014: 2; Krishnan 2009: 43). UNIDIR (2014) defines

---

[3] Bongard and Sayers (2002) further differentiate between remotely controlled, or "tele-operated", robots and self-directed, or "autonomous", robots (Bongard & Sayers 2002: 299).

[4] Depending on the domain of operation, one can distinguish between unmanned ground vehicles (UGVs), unmanned aircraft vehicles (UAVs), unmanned surface vehicles (USVs), and unmanned undersea vehicles (UUVs). The term "system" is sometimes used instead of "vehicle" as a more general definition, with acronyms UAS, UGS and UMS referring to aircraft, ground and maritime systems (DoD 2013: 2–8).

[5] AI typically refers to "narrow" or "weak" AI, which is designed for a very specific task unlike "general" or "strong" AI, which would be able to perform every cognitive task as well as or better than human (FLI 2016).

autonomy as "[...] a spectrum of capability moving from remotely controlled systems on one side to autonomous weapon systems on the other" (UNIDIR 2014: 2). However, this description may be inaccurate because, in reality, autonomy has multiple dimensions and, therefore, multiple spectrums along which it can move. Levels of autonomy can vary based on the human-machine command-and-control relationship, the sophistication of a machine's decision making, and the type of decision or function being automated (Boulanin & Verbruggen 2017: 5–7; Scharre 2015: 9–11; Scharre & Horowitz 2015: 8–9). Importantly, these three dimensions are independent. It is for this reason that an AWS can have significant autonomy in one dimension while lacking any significant autonomy in the other two (Scharre 2015: 9).

The human-machine command-and-control relationship dimension enables discerning between AWS with some human involvement, on one hand, and those with little to no involvement, on other hand (Boulanin & Verbruggen 2017: 5–6; Scharre 2015: 10; Scharre & Horowitz 2015: 6; HRW 2012: 2). Within this specific dimension, AWS can be classified as human-in-the-loop, human-on-the-loop, and human-out-of-the-loop (HRW 2012: 2). Some authors also prefer the terms semi-autonomous, human-supervised autonomous, or fully autonomous. The first one refers to systems that operate autonomously for a period of time, then stop and wait for further instructions. The second category refers to systems that operate autonomously only under supervision of a human who can intervene in case of an accident. Lastly, the third category refers to systems with no human intervention (Scharre 2015: 10; Scharre & Horowitz 2015: 6). This dimension is reflected in definition of AWS by the US DoD (see Table 1).

Next, the sophistication of the machine's decision making spectrum allows differentiating between simple AWS designs, on one hand, and complex designs, on the other (Boulanin & Verbruggen 2017: 6; Scharre 2015: 10; Scharre & Horowitz 2015: 6). Within this spectrum, AWS can be classified as automatic, automated or autonomous systems. The first one refers to systems that are merely mechanically responding to environmental or sensory input according to predefined procedures that do not account for sudden changes or uncertainties. These may include mines, trip wires or industrial robotic arms. The second type refers to systems which are more complex, operating according to sets of more general rules or accounting for some degree of environmental change. These are, for example, self-driving cars or programmable thermostats. Last type refers to complex systems that exercise high degree of self-governance, self-learning, and behavior that may be unpredictable at times. Some reserve the use of this term for intelligent entities, or AI (Boulanin & Verbruggen 2017: 6; Scharre 2015: 10; Scharre & Horowitz 2015: 6).

Table 1: US DoD Definitions of AWS

Term	Definition
Autonomous weapon system	"A weapon system that, once activated, can select and engage targets without further intervention by a human operator" (DoD 2012: 13).
Human-supervised AWS	"An autonomous weapon system that is designed to provide human operators with the ability to intervene and terminate engagements, including in the event of a weapon system failure, before unacceptable levels of damage occur" (DoD 2012: 14).
Semi-autonomous weapon system	"A weapon system that, once activated, is intended to only engage individual targets of specific target groups that have been selected by a human operator" (DoD 2012: 14).

Source: DoD (2012: 13–14)

In addition, it is necessary to specify the type of decision or functions being automated, some of which may be critical and complex while others not so much (Boulanin & Verbruggen 2017: 6–7; Horowitz 2017: 184–185; Scharre 2015: 11; Scharre & Horowitz 2015: 7). This is crucial because autonomy in functions such as navigation is relatively low-risk, and therefore "acceptable", whereas autonomy in functions such as targeting may pose significant ethical, legal, and strategic risks (Boulanin & Verbruggen 2017: 6–7). In other words: "[d]ifferent decisions have different levels of complexity and risk" (Scharre & Horowitz 2015: 7). For example, Boulanin and Verbruggen (2017) identify capability areas such as mobility, health management, targeting, intelligence, and interoperability. These involve markedly different decisions and functions (Boulanin & Verbruggen 2017: 20). Nevertheless, the controversy surrounding AWS stems prevailingly from autonomy in targeting

(Horowitz 2017: 184). For this reason, authors typically focus on autonomy "in" weapon systems rather than general level of autonomy.[6]

The continuing confusion around the definition of an "autonomous weapon system" arises precisely from these different human-, technology-, and function-centered approaches—each with its pros and cons. For instance, it is inadequate to define AWS in terms of technical parameters, typically ascribed to an object, when autonomy is more of a characteristic than a technical parameter, the differences in offensive and defensive systems are negligible, and there is excessive focus on the state-of-the-art that becomes quickly outdated. While human-centered approaches pose a challenge in establishing a threshold and means for monitoring appropriate human involvement, function-centered approaches may be so broad that they will capture many existing systems which do not necessarily constitute AWS (UNIDIR 2017b: 19–21). In addition, the problem is exacerbated by alternative labels such as "killer robots", "lethal autonomous robotics", and "lethal autonomous weapon systems" (LAWS) (see UNIDIR 2017b; Heyns 2013; HRW 2012).

Ultimately, the fundamental substance of autonomy "in" weapon systems is the removal of human operator, partially or completely. AWS represent a new form of warfare where decisions about target selection, prioritization, and perhaps even use of force, are no longer entirely under control of human operators (Krishnan 2009: 33). Despite the fact that the general level of autonomy in most systems remains relatively low, the trend is on the rise (Boulanin & Verbruggen 2017: 115; Roff 2016). Even though many systems incorporate certain autonomous functions, only few, if any, could be described as truly autonomous (Boulanin & Verbruggen 2017: 115; Horowitz 2017: 161; Scharre 2015: 12; Scharre & Horowitz 2015: 3; UNIDIR 2012: 2–3). Indeed, most of the current-day military robots such as UAVs would cluster on the lower end of all the autonomy dimensions (UNIDIR 2014: 3). Nevertheless, because AWS are understood as paradigm-shifting technology with immense military benefits, autonomous weapons are increasingly being developed by technologically advanced states (Marra & McNeil 2013: 30).

[6] In addition, ICRC definition of AWS can serve as an example of definition incorporating this dimension of autonomy. AWS is "[a]ny weapon system with autonomy in it critical functions. That is, a weapon system that can select (i.e. search for or detect, identify, track, select) and attack (i.e. use force against, neutralize, damage or destroy) targets without human intervention" (ICRC 2016: 8).

Authors often jokingly refer to today's autonomous systems as being closer to robotic vacuum cleaners and washing machines than to "Terminators" (Scharre 2015: 9; Sharkey 2010: 376). For the most part, military robots such as unmanned combat aerial vehicles (UCAVs) used in drone strikes are quite limited compared to inhabited vehicles (Horowitz 2017: 161). And though the age of robotic warfare is often associated with the advent of UCAVs in the early 2000s, in fact, systems with autonomous functions have been in use at least since the Second World War (Scharre & Horwitz 2015: 8; Krishnan 2009: 7). Nowadays, anything from guided munitions, cruise missiles, torpedoes, anti-personnel or encapsulated torpedo mines, air and missile defense systems, active protection systems, loitering weapons and robotic sentry weapons could fit the category of autonomy "in" weapon systems (Arkin 2017: 38; Boulanin & Verbruggen 2017: 36–53; Scharre & Horowitz 2015: 8–14).

As mentioned above, sophistication of autonomy in these systems remains quite low, and is relegated predominantly to mobility-related functions (Boulanin & Verbruggen 2017: 115–116). As regards the use of force, a majority of systems incorporating autonomy "in" the targeting, which includes identification, tracking, prioritization, selection, or engagement, are human-supervised defensive systems reserved for situations when the window to react is deemed too short (Altmann & Sauer 2017: 118; Boulanin & Verbruggen 2017: 115). In addition, the few offensive systems with autonomous targeting can engage only some predefined categories of targets which are determined in advance of the deployment and use by operators (Boulanin & Verbruggen 2017: 115). However, these systems do not necessarily constitute obvious examples of AWS. This is because there are no fully autonomous systems in existence as of today (Scharre & Horowitz 2015: 3–4). For this reason, it is better to think of the existing systems as "precursors" of future generation AWS (Altmann & Sauer 2017: 118, 123).

While prototypes of UCAVs with autonomy "in" functions such as take-off or landing and inflight refueling exist, for example the US X-47B (see Figure 1) or UK Taranis, vast majority of today's UCAVs are remote-controlled human-in-the-loop systems (Altmann & Sauer 2017: 122–123; Scharre & Horowitz 2015: 2). However, most commonly used human-in-the-loop systems include guided munitions such as torpedoes or cruise missiles rather than UCAVs that often come to mind in the context of today's military robotics. Guided munitions are capable of homing on a target or a location specified by the person launching that weapon (Scharre & Horowitz 2015: 8–9). Since the target is determined in advance, some authors have questioned if these systems should be perceived as AWS. In some sense, autonomy in guided munitions only supports the execution of an attack by a human operator. This technology is also relatively affordable and available on the market (Boulanin & Verbruggen 2017: 47–49).

Figure 1: X-47B (Northrop Grumman)

Source: https://goo.gl/rw9bVz

Figure 2: Phalanx CIWS (Raytheon)

Source: https://goo.gl/YtLjgZ

Human-on-the-loop systems such as air defense systems, active protection systems, shorter-range counter-rocket, artillery, mortar systems, or robotic sentry weapons constitute a more accurate example of AWS precursors (Boulanin & Verbruggen 2017: 36–47; Scharre & Horowitz 2015: 12–13). All of these are human-supervised defensive systems used to protect bases, vehicles and borders from pre-programmed targets. With the exception of robotic sentry weapons, they are all anti-materiel weapons. Additionally, human operators supervise the system and can intervene when necessary (Scharre & Horowitz 2015: 12–13). Air defense systems such as the US Phalanx CIWS (see Figure 2) are designed to defend areas around ships—on which they are stationed—by detecting, tracking and engaging incoming air projectiles. Robotic sentries such as the Korean Super aEgis II (see

Figure 3), in comparison, can be stationary or mounted to a vehicle, and they are anti-personnel, which is relatively rare (Boulanin & Verbruggen 2017: 36, 44).[7]

Finally, loitering weapons constitute the only example of human-out-of-the-loop systems at this moment (Boulanin & Verbruggen 2017: 50–54; Scharre & Horowitz 2015: 13–14).[8] Also known as "suicide drones", loitering weapons fit a category between guided munitions and UCAVs. Though most of these systems are remote-controlled, some engage targets autonomously. The crucial difference from the aforementioned systems is that loitering weapons are not aimed at a particular target but rather a general class of targets such as ships, radars, and tanks, within a deployment area. Once the weapon detects a target that matches the class, it dives onto it. Thus, unlike other mentioned weapons, it is designed to conduct offensive missions as well as defensive (Boulanin & Verbruggen 2017: 50–53; Scharre & Horowitz 2015: 13). The Israeli Harpy or Harop system (see Figure 4) is currently the only example of anti-radar loitering weapon with full autonomy in engagement (Scharre & Horowitz 2015: 13).

Figure 3: Super aEgis II (DoDAAM)

Source: https://goo.gl/UtTefo

Figure 4: Harop (IAI)

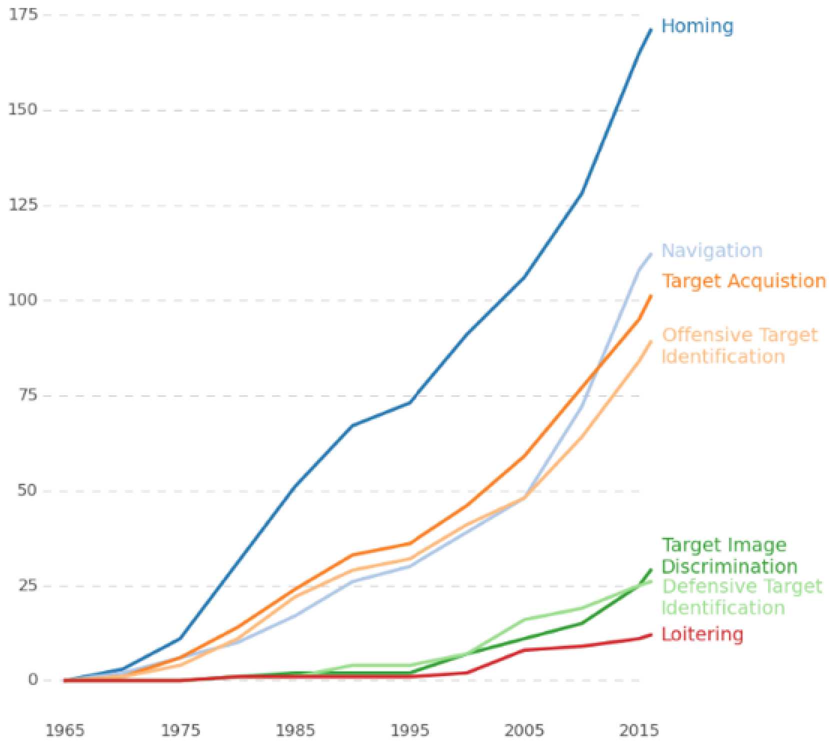
Source: https://goo.gl/kWb9kK

[7] Israel and South Korea are the two known developers of robotic sentries (Boulanin & Verbruggen 2017: 44)

[8] Scharre and Horowitz also mention encapsulated torpedo mines as an example of human-out-of-the-loop systems. It should be noted that mines are generally omitted from discussions on AWS since they are considered very simple, predictable devices, in case of which the person deploying the weapon knows the location or the type of target in advance. The problem with encapsulated torpedo mines is that instead of exploding once triggered, the device releases a torpedo, which then itself autonomously seeks the target. At that point, the person launching the weapon is "out of the loop". The Russian PMK-2 constitutes the only known system in use today (Scharre & Horowitz 2015: 14–15).

Before turning to individual drivers of the increasing autonomy, it is worth exploring some of the recent trends in AWS development. Generally, surveys show that weapon systems are becoming steadily more autonomous (see Figure 5). Roff (2016) finds that there has been a push in the top five weapons exporting counties for greater autonomy, particularly in target identification and learning functions (Roff 2016). According to some estimates, defense spending on military robotics will reach 7.5 billion USD by the end of 2018—a dramatic increase compared to previous years (Horowitz 2017: 163). Investments in research and development (R&D) of military robotics grow rapidly in the US as well as other states (Boulanin & Verbruggen 2017: 94–103; Horowitz 2017: 165). The US identified robotics as a main enabler of maintaining military dominance within its "Third Offset" strategy, Russia intends to replace 30% of all military technology with robots by 2025, and China seeks to become a leader in AI by 2030 (ILW 2017: 1–2; Kania 2017: 4–6).

Figure 5: Deployed Weapons Systems with Autonomous Functions (1965–2015)[9]



Source: Roff (2016)

[9] The data covers 284 systems from the US, China, Russia, Germany and France (Roff 2016).

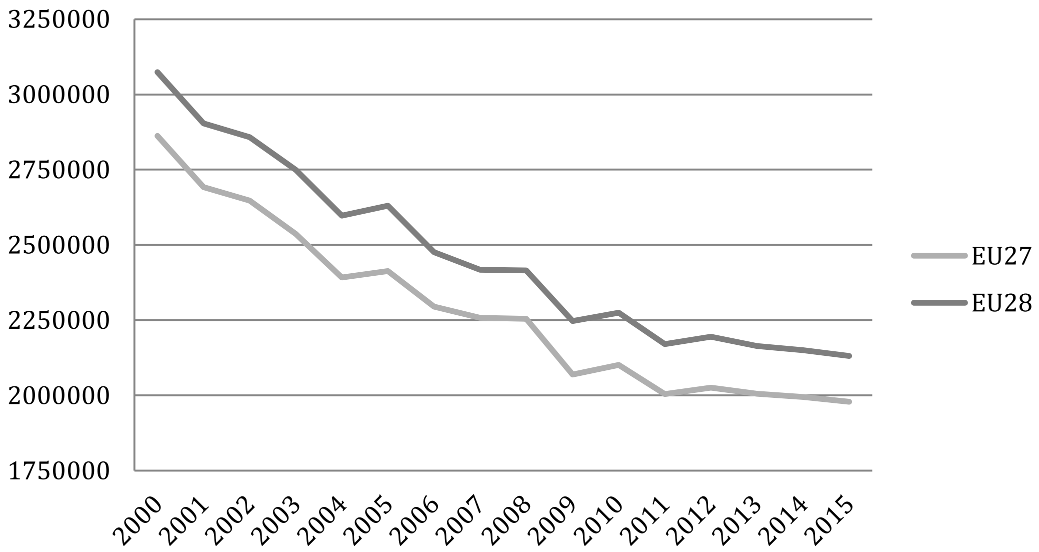
2.2. Drivers of Autonomy: The Rationale behind AWS Development

With regard to the previously mentioned trends, it is clear that there is now a global push for increasing the autonomy in weapon systems. Logically, autonomy must hold immense utility for military and national security if countries are willing to spend steadily larger portions of defense budgets on it. But what exactly are the benefits that drive this push? This section seeks to elaborate on some of the crucial drivers of autonomy such as force substitution, costs reduction, suitability for dull, dirty, and dangerous operations, casualty reduction, speed, accuracy, precision, or the potential for improving ethical conduct on the battlefield, and developments in the civilian sector (See Arkin 2017; Boulanin & Verbruggen 2017; Etzioni & Etzioni 2017; UNIDIR 2017b; Marchant et al. 2011). These are some of the most frequently mentioned benefits and thus drivers of increasing autonomy.

Force substitution. First, one of the crucial reasons for increasing autonomy is the concern over decreasing manpower pool for military recruitment (Marchant et al. 2011: 275; Krishnan 2009: 2, 35–36). Certain tasks can be delegated to AWS so that fewer soldiers are needed, and the efficiency of each soldier increases (Arkin 2017: 36; Etzioni & Etzioni 2017: 72). This benefit is particularly relevant for technologically advanced countries, where issues of recruitment are prevalent as fewer people are fit for military service each year. The situation will likely only deteriorate, forcing those countries to substitute soldiers with robots over time. In addition, it will probably become increasingly difficult to sustain current defense spending in some states. In this context, incorporating AWS will enable reducing personnel-related costs (Krishnan 2009: 35–36). For illustration, Figure 6 on the following page shows decline in total military personnel in the EU.

Costs reduction. A closely related driver of increasing autonomy is reducing the costs of operating weapon systems (Boulanin & Verbruggen 2017: viii; Etzioni & Etzioni 2017: 22; Krishnan 2009: 2). According to estimates, a single soldier in Afghanistan costs the US some 850,000 USD annually compared to some 230,000 USD annually for a weaponized TALON UGV. (Etzioni & Etzioni 2017: 72). Robotic platforms cost far less than manned platforms, and they are substantially cheaper over their lifetime compared to soldiers. Indeed, predictions suggest that a robot would cost only 10% of a solider (Krishnan 2009: 2). By the same token, improving the efficacy of soldiers by introducing AWS would bring major long-term savings (Boulanin & Verbruggen 2017: viii). This possibility in combination with dwindling budgets presents a significant driver of increasing autonomy (Krishnan 2009: 2).

Figure 6: Military Personnel, Total (2000–2015)[10]



Source: Author (2018)

Dull, dirty, and dangerous missions. Substituting soldiers with robots can be particularly useful in dull, dirty, and dangerous missions (UNIDIR 2014: 6; Lin et al. 2008: 6; DoD 2007: 19). Certain tasks which are dull, for example, extended reconnaissance missions, usually require high degrees of concentration over long period of time. However, soldiers' concentration drops rapidly over time (Krishnan 2009: 40; Lin et al. 2008: 6). Military robots, on the other hand, never get tired or distracted. They are also more suitable for dirty missions that potentially involve sample collection after the nuclear or biochemical attack. By substituting soldiers with robots, military can also decrease the risk of soldiers' exposure to hazardous substances. Lastly, robots are much more suitable for dangerous missions such as explosive ordnance disposal. In this context, substituting soldiers with robots

[10] Data come from the World Bank website (available at: https://data.worldbank.org/).

reduces the political and human cost of mission failure (UNIDIR 2014: 6; Lin et al. 2008: 5; DoD 2007: 19).

Casualty reduction. The potential of reducing one's own casualties presents another important driver (Arkin 2017: 37; US Army 2017: 2; Marchant et al. 2011: 275). In some cases, soldiers may be removed from mission completely, keeping them from harm's way (Arkin 2017: 37; Etzioni & Etzioni 2017: 72). In other cases, robots can be used to protect soldiers by providing them with greater situational awareness and standoff distance, thereby reducing their exposure and improving their survivability (US Army 2017: 1–2; UNIDIR 2014: 6). This goes hand in hand with two crucial tactical advantages: extending warfighter's reach and expanding the battlespace. Robots allow combat to take place in very remote areas, and they enable soldiers to reach these areas, striking farther (Arkin 2017: 37; Marchant et al. 2011: 275).

Speed. Another important driver of increasing autonomy is the speed with which robots can conduct their tasks (Boulanin & Verbruggen 2017: vii; US Army 2017: 1; UNIDIR 2014: 6; Krishnan 2009: 40). For example, lowering response times and creating instant effects on the battlefield is essential in achieving victory (Krishnan 2009: 40). Robots can respond much faster than humanely possible both in the physical and cognitive domain (US Army 2017: 1; Etzioni & Etzioni 2017: 72). In case of the latter, robots can process vast amounts of information, and collect and prioritize data, thereby facilitating decision-making (US Army 2017: 1). Moreover, it may be necessary to increase autonomy in decision-making in future since "[...] military systems (including weapons) [...] will be too fast, too small, too numerous and will create an environment too complex for humans to direct" (Adams 2002: 2). Robotic systems will also achieve all the above mentioned with greater precision and accuracy (UNIDIR 2017b: 13–14).

Improving ethical conduct. Further driver of increasing autonomy concerns the potential of robots to improve ethical conduct on the battlefield (Etzioni & Etzioni 2017: 74; UNIDIR 2014: 6; Arkin 2010: 333; Lin et al. 2008: 6). Specifically, robots could police ethical behavior of soldiers by monitoring their conduct and by reporting infractions (Arkin 2010: 333–334). There is, however, a more daring proposition—that is, they could act more ethically than soldiers (see Arkin 2010).

Unlike soldiers, robots are unaffected by negative emotions, which often result in criminal behavior (Etzioni & Etzioni 2017: 74; Arkin 2010: 333; Lin et al. 2008: 6). In addition, robots need not be programmed with self-preservation, while they can be programmed to respect international humanitarian law (IHL) and human rights law (IHRL), which could reduce breaches thereof (UNIDIR 2014: 6; Arkin 2010: 333). Substituting soldiers with robots could therefore reduce the number of war atrocities (Lin et al. 2008: 6–7).

Civilian sector. To a large extent, the drive for increasing autonomy can be attributed to the civilian sector (NDU 2017: 4–6; UNIDIR 2014: 6–7). After all, it is the innovation within the civilian sector which makes many military applications possible. Advances in autonomy are possible precisely due to the expectations and investment of consumers and the industry. Furthermore, as civilian applications become more commonplace, public perception of AWS may become increasingly positive, potentially creating a demand for AWS R&D and use (UNIDIR 2014: 6–7). These trends are supported by estimates according to which market for mobile civilian and military robotics will exceed 14 billion USD by 2019. Large companies such as Google and Facebook have invested heavily in machine learning and AI development (NDU 2017: 4–6). Accessibility and availability combined with low prices of components thus become key enablers for increasing autonomy.

Other drivers. Some other less-known drivers should be briefly mentioned. For example, unlike remote-controlled systems, AWS do not depend on secure data links, which are used to transmit control and sensory data to a control station. The reliance on secure data links presents a challenge because cables can be physically destroyed, satellites have limited bandwidth, and wireless data can be disrupted (Krishnan 2009: 37–38). In addition, with regard to logistics and maneuverability, AWS seem to provide major benefits as well. Robots provide army logistics with an option to move resources to the most forward tactical points without the risks of soldiers' exposure. Due to the speed of information processing, they will further be able to outmaneuver enemies' systems, thus expanding the time and space within which armed forces can operate (US Army 2017: 1–2).

2.3. Perils of Autonomy: Ethical, Legal, Security, and Operational Challenges

Benefits usually come with costs, and so do the benefits of AWS. In fact, the literature on AWS seems to be oversaturated with images of doomsday devices, or Terminators. However, beyond this oversimplification, what exactly are the perils of increasing autonomy? This section elaborates on various ethical, legal, security, and operational challenges posed by AWS. In particular, it focuses on issues such as arms-racing, proliferation, strategic instability, complexity, unpredictability and unreliability, vulnerability to cyber-attacks, inability to uphold principles of IHRL and IHL, violation of human dignity, potential of public backlash, risk-free warfare, and certain others (see Altmann & Sauer 2017; UNIDIR 2017b; Crootof 2016a; Scharre 2016; Heyns 2013; Asaro 2012; HRW 2012; Sharkey 2012b; Sparrow 2007). These are the most frequently mentioned perils of autonomy in weapon systems.

Arms-racing. Many authors argue that the reckless pursuit of AWS R&D has triggered an arms race between major powers, including the US, Russia or China (Boulanin & Verbruggen 2017: viii; FLI 2017; Geist 2016: 318; Sharkey 2012a: 798). The concern is that, the push for greater autonomy on part of major powers could trigger action-reaction cycle with the result of decreased security for all (Dhanapala 2017: 57; FLI 2015; Altmann 2009: 71). To some extent, the process can be observed in the field of remote-controlled UCAVs (Fuhrmann & Horowitz 2017a: 397; Altmann 2013: 141; 2009: 76). One of the vital differences between AWS and other fields of military technology is that much of the innovation comes from the commercial sector, and this can contribute to a faster pace of arms-racing (Scharre 2015: 5). Furthermore, the 2015 open letter by the Future of Life Institute signed by renowned scientists warns that a global arms race would be "virtually inevitable" were the military powers to double down on AWS R&D (FLI 2015).

Proliferation. A related issue concerns the proliferation of AWS (Altmann & Sauer 2017: 117–118,; Dhanapala 2017: 57–61; Scharre 2015: 5–6; Altmann 2013: 137, 141; Sharkey 2012b: 122–123; Altmann 2009: 78; Lin et al. 2008: 82). The main difference between AWS and other paradigm-shifting technologies is that AWS can be relatively easily obtained, and this is because the foundations of the technology lie in its software, which can be copied effortlessly (Altmann & Sauer 2017: 125–126). Furthermore, as much of the innovation comes from the civilian sector, the technology will become increasingly available, including on grey and black markets, to state and "non-state" actors (Altmann & Sauer 2017: 117, 127; Scharre 2015: 5–6).[11] This becomes a challenge especially if AWS fall into hands of rogue states or terrorist organizations (Altmann 2013: 137, 141; Dhanapala 2017: 57; Sharkey 2012b: 122–123). The current day UAV proliferation can be seen as a precursor to AWS proliferation. In fact, more than 90 state and non-state actors already possess advanced drones (Ewers et al. 2017: 8). It can be expected that the AWS will follow the same trajectory as today's UAVs.[12]

Strategic instability. Both arms-racing and proliferation are further linked to strategic instability (Altmann & Sauer 2017: 119–121, 128–130; Dhanapala 2017: 57; Scharre 2017: 15, 20–26; Altmann 2013: 141; 2009: 76–77). Over the course of history, weapons threatening the maintenance of status quo have been seen as undermining stability, thereby increasing the likelihood of war (Scharre 2017: 19–20). AWS could be destabilizing because the speed they provide creates incentives to strike first so as to gain significant advantage or outright destroy the opponent (Altmann & Sauer 2017: 119; Scharre 2015: 15, 20). Incentives to strike first would be further exacerbated by the fact that mission failure would not result in direct casualties, and since AWS can be equipped with stealth and swarming capabilities or nuclear weapons (Altmann & Sauer 2017: 130; Altmann 2013: 141; 2009: 76–77). For these reasons, deployment of AWS would destabilize global or regional security, and this could lead to "flash wars" (Dhanapala 2017: 57; Scharre 2015: 20).

C*omplexity*. In complex and tightly coupled systems such as AWS, accidents and hidden interactions are inevitable (Borrie 2017: 23; Wallach 2017a: 297; Scharre 2016: 7, 25–37; UNIDIR 2016: 3, 6–10; Lin et al. 2008: 8). In any such

---

[11] Similarly to current-day drones, autonomous systems will be largely based on dual-use technologies, and will thus exacerbate the issues of regulation, proliferation, and so on (see UNIDIR 2017a). Gradual introduction of AWS into society can be expected even at times of peace, which can have a negative impact on privacy and civil security (Lin 2010: 323; Lin et al. 2008: 83–84).

[12] In addition, Fuhrmann and Horowitz (2017a) find that most influential factors explaining why states seek to obtain drone technology include state's technological capacities, state's involvement in territorial disputes and counterterrorism. Interestingly, both democracies and autocratic regimes are more likely to obtain drone technology compared to mixed regimes (Fuhrmann & Horowitz 2017a: 397).

systems, on occasion, the so-called normal accidents and black swans occur in spite of any wrongdoing on part of operators (Wallach 2017b: 31; UNIDIR 2016: 10; Sagan 2013: 13). Consequently, no AWS can be designed fail-safe, regardless of good intentions of its designers and engineers (Borrie 2017: 23; UNIDIR 2016: 3). Furthermore, adding more safety features can only exacerbate the issue because the system would become even more complex (UNIDIR 2016: 11). The additional complexity would also require thorough training of personnel in order to improve their ability to understand, predict, and manage AWS interactions. (Hoffman et al. 2016: 248, 251; Scharre 2016: 7). Finally, the complexity scales with the functional complexity and lack of human oversight (Scharre 2016: 7, 11).

Unpredictability and unreliability. As a result of the complexity, AWS can be unpredictable and unreliable (Wallach 2017a: 295–298, 308–310; 2017b: 31–32; Scharre 2016: 7–11, 18; UNIDIR 2016: 8). There is a significant risk in regard to probability that AWS will not behave as intended, and this may lead to unintended engagement and fratricide, severe civilian casualties, or crisis escalation (Wallach 2017a: 298; 2017b: 31; Scharre 2016: 18). While many actors, including even the poor states and non-state actors, will seek to obtain the technology, adequate tests, maintenance, and personnel training will be difficult and costly even for advanced states (Wallach 2017a: 309–310). Regardless of capacities, there are currently still no adequate verification methods to produce proofs of behavior of machine-learning-based systems (Wallach 2017b: 31–32; UNIDIR 2016: 8). The degree of acceptable risk will further vary across countries, and some may seek to obtain them despite high risks (Wallach 2017a: 308). Table 2 summarizes some examples of reported robotic weapons accidents (see Table 2).

Vulnerability. Malfunctions may also occur as a consequence of adversarial efforts (Wallach 2017a: 303; UNIDIR 2017a: 10–12; 2016: 11; Scharre 2016: 8, 36). As with any software-based systems, AWS would be susceptible to hacking, particularly through zero-day exploits that are very difficult to detect in complex systems (UNIDIR 2017a: 10–12). Furthermore, techniques such as jamming and electromagnetic pulses can be used to disrupt the communications links (UNIDIR 2017a: 11; UNIDIR 2016: 11). In addition, with the view to three previous points, the problem of trust should be mentioned (Roff & Danks 2017: 3–4; NDU 2017: 5;

Scharre 2015: 5). Roff and Danks (2017) argue that deployment of new military technologies requires not only predictability or reliability but also trust. However, when the machine is out of human control—that is, fully autonomous—it cannot be trusted (Roff & Danks 2017: 3–4). Aside from these operational issues, technical shortcomings can also impact compliance with IHL and IHRL.

Table 2: Reported Accidents Involving Robotic Weapons[13]

System	Description of events
Patriot missile	In March 2003, Iraq, US deployed Patriot missile engaged friendly aircraft, causing a fratricide while confusing the aircraft with enemy missiles (Hoffman et al. 2016: 248).
Oerlikon GDF-005 anti-aircraft cannon	In October 2007, South Africa, a semi-autonomous cannon deployed by the South African army malfunctioned, killing nine soldiers and wounding 14 others (Shachtman 2007).
TALON SWORDS UGV	In April 2008, Iraq, US deployed UGVs were grounded after an incident when the robots started to move without being commanded, allegedly pointing guns at friendly soldiers (Lin et al. 2008: 7).
MQ-1 Predator UAV	In February 2010, Afghanistan, US drone crew mistakenly identified civilian vehicles as an insurgent convoy killing or injuring of at least 35 civilians (Hoffman et al. 2016: 248).

Source: Author (2018)

Indiscriminate attacks. Many authors point to the potential inability of AWS to comply with principles of IHL, particularly the principle of distinction (Crootof 2016a: 1873–1874; Cass 2015: 1035; Heyns 2013: 13; HRW 2012: 30–32; Sharkey 2012b: 116–120; Krishnan 2009: 99). Today, there are no sensors, which would enable distinguishing between combatants and non-combatants. Even though the current technology allows face recognition or distinguishing between human and non-human, it cannot identify "hors de combat", and individuals in the process of surrendering, or differentiate between civilians and insurgents (Cass 2015: 1035; Heyns 2013: 13; Sharkey 2012b: 116). This is doubly problematic in complex environments such as urban warfare, where terrorists or insurgents can trick AWS by disguising as civilians and by exploiting their low sensory capabilities to target

[13] Accidents selected here are illustrative only.

civilians instead (Geiss 2015: 14; Krishnan 2009: 99). The issue is that assessment of individual's status depends on the context, interpretation of which is necessarily subjective and requires human judgment (Crootof 2016a: 1873).[14]

Disproportionate attacks. Aside from the aforementioned, fully autonomous would likely be unable to comply with the proportionality principle of IHL (Crootof 2016a: 1876–1881; Cass 2015: 1037; Heyns 2013: 13–14; HRW 2012: 32–33; Sharkey 2012b: 123; 2010: 380). AWS of today cannot sense and calculate whether military advantage gained from attack would outweigh collateral damage. AWS use might therefore lead to civilian casualties, unnecessary suffering and superfluous injury (Crootof 2016a: 1876; Cass 2015: 1037; HRW 2012: 32; Sharkey 2012b: 123; 2010: 380). Similarly to the principle of distinction, the role of the context matters greatly because permissible collateral damage depends on it. Furthermore, human judgement and good faith are necessary for adequate calculations as well (Crootof 2016a: 1876; Heyns 2013: 13–14). As regards the principles of IHL, it is therefore questionable, whether AWS will ever have the capacity to imitate human decision-making involved in such calculations (HRW 2012: 33).

Accountability gaps. The use of AWS could create accountability gaps when a breach of law occurs (Arnold 2017: 14–15; Davison 2017: 16–17; Cass 2015; Heyns 2013: 14; HRW 2012: 42–44; Sharkey 2012b: 117; Sparrow 2007: 69–73). More precisely, it is unreasonable to expect that robots could be held criminally liable for faulty decisions or malfunctions. Robots are not individuals who can be guilty of a crime, and sanctioning them will have no real effect (Arnold 2017: 14; Sharkey 2012b: 117). Although other subjects can be held accountable, attribution will still present a significant challenge in case of all subjects. Commanders cannot reprogram the system in real time once it malfunctions; manufacturers can avoid liability by disclosing risk of malfunctions to military; the number of programmers involved in coding can make identifying a single offender impossible; and states are barely deterred by a threat of reparation to civilian victims (Cass 2015: 1049–1053; HRW 1012: 43–44; Sparrow 2007: 69–73).[15]

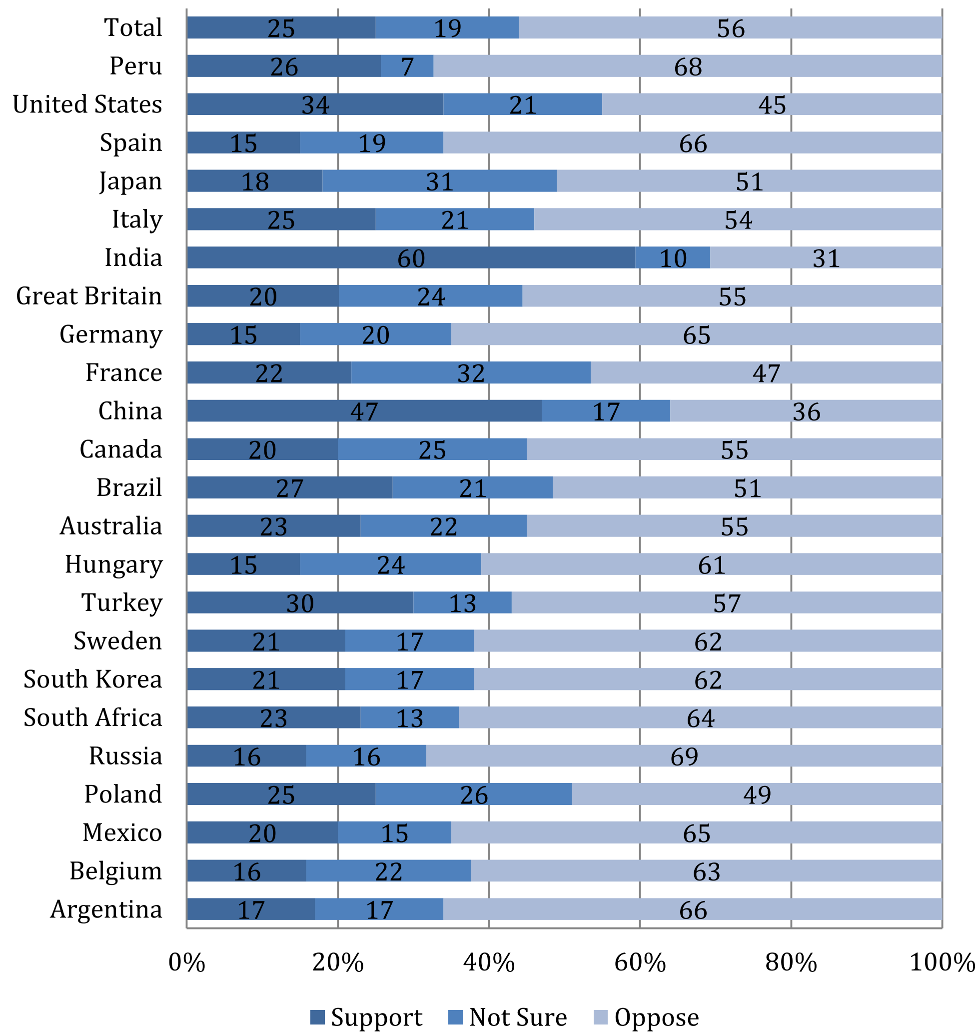
[14] Moreover, because of the absence of a clear definition of a combatant, AWS cannot be preprogrammed with such (Cass 2015: 1035).

[15] In addition, authors argue that AWS cannot be held accountable because they lack full moral agency (Heyns 2013: 14; Asaro 2012: 693).

Article 36 weapons reviews. Another challenge relates to ensuring that AWS development respects principles of IHL as mandated by the legal weapons reviews procedure described by the Article 36 of the Additional Protocol I to the Geneva Conventions (Davison 2017: 10; Ford 2017: 82–83; Asaro 2012: 692–693; Sharkey 2012b: 119). Some countries such as the US are not party to the protocol, and even though this process is considered an international custom, they may circumvent the requirement by combining existing technologies. This was the case with drone technology and hellfire missiles (Cass 2015: 1040–1041; Sharkey 2012: 119). It is also unclear how complex learning systems such as AWS can be tested or verified so as to ensure predictability and reliability, let alone compliance with IHL or IHRL, and how to translate results of these tests into informed legal policies (Davison 2017: 10; UNIDIR 2017b: 4; Ford 2017: 82).

Violation of human dignity. Perhaps the most controversial issue relates to ceding the decision over life and death to a machine (Mayer 2015: 77; Heyns 2013: 17; HRW: 2012: 33, 37–38; Sharkey 2012b: 116–118; Wallach & Allen 2012: 127; Krishnan 2009: 130; Sparrow 2007: 65–67). Former Special Rapporteur to the UN, Christof Heyns suggested that "[m]achines lack morality and should as a result not have life and death powers over humans" (Heyns 2013: 17). It is very questionable whether AWS should be entrusted with lethal decision-making, when they lack full moral agency, or have no understanding of the value of human life (Geis 2015: 18; Wallach & Allen 2012: 693). Even in certain situations when taking human life is in accordance with IHL, the decision is always preceded by deep moral deliberation, including the ability to feel compassion or show mercy. However, AWS cannot feel empathy or relate to individuals (Geiss 2015: 18; HRW 2012: 33; Sharkey 2012b: 116–118). As a result, AWS lack the natural inhibition of humans to kill (Krishnan 2009: 130).

Figure 7: Opinion Poll on Use of AWS, 2017



Source: Ipsos (2017)

Public backlash. AWS development may also lead to a public backlash (Ipsos 2017; ORi 2015; UNIDIR 2015: 6; ICRC 2014: 92; Roff 2017; Lin et al. 2008: 9). In particular, it is questionable whether AWS use would be contrary to the principles of humanity and dictates of public conscience found in the Martens clause.[16] It is the dictates of public conscience, which should decide on the acceptability of AWS in the absence of rules specifically aimed at the technology (Roff 2017; ICRC 2014: 92). Thus, when norms of public conscience emerge in the form of public opinion,

[16] Principles of humanity and dictates of public conscience are considered customary international law. This international custom was first formalized in the Martens clause, which can be found in the Preamble of the Hague Conventions II of 1899 and IV of 1907 (ICRC 2014: 92).

that go contrary to investment into AWS R&D, the sense of moral obligation should incentivize countries to refrain from such course of action (UNIDIR 2015: 6). When ignored, this could lead to a public backlash (Roff 2017; Lin et al. 2008: 9). Early surveys seem to suggest that public sees AWS as problematic (see Figure 7) (see Ipsos 2017; Horowitz 2016; ORi 2015).

"Risk-free" warfare. Authors suggest that AWS development and deployment could lower the threshold, or political costs, of declaring a war (Dhanapala 2017: 57; Asaro 2012: 692; Sharkey 2012b: 122; Lin 2010: 313, 321–324; Krishnan 2009: 2). By removing some risks associated with war, in particular by reducing one's own casualties, AWS use could make war a preferred and even convenient method of dispute settlement. It is the perceived lack of risk, or the lack of "body bags" returning back home, and thus lower social, economic, and political costs of war, which could exacerbate a culture of war (Lin 2010: 313, 321; Sharkey 2012b: 122). Aside from increasing the likelihood of war, risk-free perception could also escalate the pace of warfare once it occurs (Dhanapala 2017: 57). Obviously, in reality, there is no such thing as a war without risks but misperceptions could have dire consequences.

Asymmetrical warfare. Nations developing and deploying AWS would likely have to face more as well as more severe asymmetrical threats (Mayer 2015: 76; UNIDIR 2015: 9–10; Asaro 2012: 692; Sharkey 2012b: 122; Lin 2010: 326; Lin et al. 2008: 80–81). In spite of the fact that the use of advanced technologies against inferior opponents is fairly common throughout history, it can be expected that the effects of AWS could exacerbate the issue of asymmetrical warfare (UNIDIR 2015: 9–10). The use of AWS could shift opponents' counter-strategies rapidly toward causing civilian casualties as attacking AWS would not have the desired effect of inflicting military casualties (Mayer 2015: 76). It is thus conceivable that opposing forces would engage in acts of terror in response (Lin et al. 2008: 81). In addition, it would be difficult to win hearts and minds of the people in counterinsurgency operations since military robots are often perceived as cowardly and dishonorable (Sharkey 2012b: 122; Lin 2010: 326; Lin et al. 2008: 80).

Other concerns. Authors further mention a variety of other challenges such as "the strategic robot problem". If militaries entrust robots with making strategic, tactical and operational decisions, such as in case of targeting procedures, this will undermine existing command and control structures, eliminating human influence over the trajectory of war (Roff 2014: 212). Additionally, the use of AWS could lead to devaluation of military service. The ideal of soldier or warrior has always been synonymous with values such as heroism, honor, courage and selflessness, and the act of taking human life has been socially permissible by upholding of these values. Thus, entrusting AWS with decisions to kill could decrease the value of soldiers' sacrifice (Mayer 2015: 75; UNIDIR 2015: 10–11).[17] Furthermore, operators could experience moral disengagement because AWS will further increase the distance between the attacker and its target, while decreasing the fear of being killed and resistance to killing (Sharkey 2012b: 111; Sharkey 2010: 371–372).[18]

[17] The conflict between the ideal of military service and military robotics is evidence, for example, by the 2014 decision by US Defense Secretary Hagel not to issue a medal to honor drone pilots after the protests within military (UNIDIR 2015: 10–11).

[18] This phenomenon is often described in connotations of "video-game" mentality (Sharkey 2012: 113). Pilots have reportedly said about drone strikes: "[i]t's like a video game. It can get a little bloodthirsty. But it's fucking cool" (Singer 2009: 308–309), or "[k]illing people is like squashing an ant. I mean, you kill somebody and it's like 'All tight, let's go get some pizza'" (Singer 2009: 391–392).

3. Literature Review

The purpose of this chapter is to provide a review of the empirical literature on AWS, particularly with a focus on different approaches and regulatory solutions to AWS found in the policy or legal literature. The chapter points to various gaps in the literature, as well as certain oversimplifications and contradictions. The author suggests that the existing literature on AWS can be divided into four approaches—a legal approach that perceives AWS mainly from the perspective of compliance or application of IHL (see Crootof 2016a; 2015; Anderson & Waxmann 2013; Schmitt & Thurnher 2013), an ethical approach which explores issues of responsibility, or moral agency, and the act of entrusting machines with kill decisions (see Wallach 2017b; Sharkey 2012a; Asaro 2012; Sparrow 2007), a technical approach that sees the issue of AWS primarily in the context of issues of reliability, predictability and various technical improvements (see Arkin 2017; Theodorou et al. 2017; Wallach 2017a; Keeley 2015; Arkin et al. 2009), and security approach, focusing on certain implications of AWS to international security, strategic stability, proliferation and arms races (see Altmann & Sauer 2017; Lachow 2017; Rickli 2017; Scharre 2017). The chapter also provides an overview of all regulatory solutions proposed so far, including the likes of a full ban, moratorium, conventions, codes of conduct, and so on (see Dunkelberg 2016; Crootof 2106a; Geiss 2015; Marchant et al. 2011). Latter part of the chapter then allows the author to establish a base for spatial analysis in constructing a policy continuum based on the proposed regulatory solutions.

3.1. Perspectives on AWS: Between “Roboutopia” and “Robocalypse”

Literature on AWS and socio-political challenges of emerging technologies in general has grown immensely over the last few years. Unfortunately, the debate suffers from a lack of theoretical framing, which is often reduced to lists of reasons for and against AWS R&D and deployment.[19] Paul Scharre (2016) notes that there

[19] In fact, articles such as “Pros and Cons of Autonomous Weapons Systems” by Amitai and Oren Etzioni are quite symptomatic of this trend (see Etzioni & Etzioni 2017).

are two basic intuitions which implicitly shape the discourse on AWS—one can be characterized as "roboutopia" and the other as "robocalypse" (Scharre 2016: 6). The optimists vs. pessimists divide can be observed across most contributions on AWS. Authors often use arguments from different theories irrespective of their background. This often creates a bias towards one side of the divide. For example, authors such as Noel Sharkey or Ronald Arkin, although both roboticists, hold very different views. One argues that use of AWS is unethical while the other argues AWS could be more ethical than soldiers (Sharkey 2012a: 797; Arkin 2010: 336).

With that in mind, authors often come to same conclusions not necessarily because of particular theoretical approach they use but rather because of optimist or pessimist attitude. Therefore, assessment of implications of AWS development differs within approaches. Of these, we can name at least ethical, legal, security, or operational approaches covered in the previous chapter. Regardless, as previously mentioned, authors tend to borrow arguments, and thus no sharp boundaries exist between approaches to AWS. Ethicists and philosophers justify their claims against AWS on the supposed inability of the technology to comply with principles of IHL, whereas roboticists and computer scientists suggest that AWS can improve ethical conduct and compliance with IHL (see Sharkey 2012a; Arkin 2010; Sparrow 2007). When advocating for regulation of AWS, legal scholars refer to security risks such as proliferation and arms-racing, and scholars of International Relations often refer to operational risks (see Altmann & Sauer 2017; Lachow 2017; Guiora 2017; Crootof 2016a).

Literature by scholars of international law is by far the most voluminous. In their contributions, authors usually focus on the broader Just War Theory or some particular concept of international law such as principles of IHL, accountability or criminal liability, current or potential forms of regulation (see Crootof 2016a; Cass 2015). On one hand, authors such as Anderson and Waxman (2017), and Schmitt and Thurnher (2013) maintain that AWS are not inherently unjust, compliance with IHL is possible, and therefore no extra regulation besides national regulation is necessary (Anderson & Waxman 2017: 1111–1112; 2013: 21–22; Schmitt & Thurnher 2013: 279). On the other hand, authors such as Vivek Sehrawat (2017), Rebecca Crootof (2016c), Hadji-Javen and Hristovski (2017) suggest that current prohibitions on unlawful use of weapons may be inefficient, and some form of regulation of AWS may be needed to uphold IHL, whether a new protocol, treaty, or a norm (Hadji-Janev & Hristovski 2017: 336–339; Sehrawat 2017: 39–40; Crootof 2016c: 1402). Various legal scholars also focus on the idea of "meaningful human control" as a standard (see Chengeta 2017; Rosert 2017; Crootof 2016b; Neslage 2016; Roff & Moyes 2016; Horowitz & Scharre 2015).

Generally, legal contributions on AWS explore adequacy and applicability of IHL to AWS. The prevailing conclusions are that AWS are not illegal per se, and that they are not unregulated, but adoption of regulation may be necessary in future so as to clarify the conditions of legal use. In contrast, ethicists or philosophers, most of which draw on the Just War Theory or particular concepts such as moral agency and responsibility, question moral acceptability rather than legal permissibility.[20] It is the issue of leaving the decision over life and death to AWS which lies at heart of ethical approaches. Noel Sharkey (2012a) and Peter Asaro (2012) note that this would constitute an affront to human dignity since AWS cannot understand human emotions, the context, and other intangible barriers to violence (Sharkey 2017: 181; 2012a: 799; Asaro 2012: 708–709). Robert Sparrow (2007) emphasizes the paradox of entrusting AWS with autonomy, implying responsibility, for which they cannot be held morally accountable (Sparrow 2007: 72). In contrast, others argue that if AWS could improve ethical conduct, their use would be morally obligatory (Wallach 2017b: 29).

This view is shared mainly by roboticists, physicists, computer scientists, and the like, who focus on technical solutions to problems posed by AWS. Without a doubt, the most advanced research on this topic has been conducted by Ronald C. Arkin (2017), who is one of the authors of the "ethical governor" software. The sole purpose of this software is to help AWS in assessing the permissibility of the use of lethal force under given circumstances, ensuring compliance with IHL (Arkin et al. 2009: 1). Thomas Keeley (2015) presents safety design features of "auditability" and "traceability" to allow detecting or fixing problematic behavior (Keeley 2015: 196). Theodorou et al. (2017) suggest transparent designs for AWS to allow real-

[20] UNIDIR 2015: 5. Authors employ both consequential and deontological ethical approaches. While the former finds moral justification in the consequences of an act, the latter finds moral justification in the nature of an act itself (UNIDIR 2015: 3).

time inspections of suspect behavior (Theodorou et al. 2017: 236–237). A bulk of literature also concerns issues of reliability or predictability, drawing on complex systems theory (see Hagström 2017; Wallach 2017a; Scharre 2016; UNIDIR 2016). Others focus on issues of human-machine interaction, concepts such as OODA loop, and so on, based on policy/doctrinal frameworks (see UNIDIR 2017a; Scharre & Horowitz 2015; DoD 2013; 2012; 2007; Marra & McNeil 2013).

Finally, in International Relations literature, the topic of AWS and emerging technologies in general comes up quite rarely. When it does, however, it is often in the context of great power competition, global security, and changing international norms. Authors have relied primarily on theories of arms control, and concepts such as arms-racing, proliferation, and strategic stability, to examine the issues (see Altmann & Sauer 2017; Lachow 2017; Rickli 2017; Scharre 2017). Authors such Jürgen Altmann (2013) suggest that AWS, or, more specifically, UVs, require preventive arms control measures, and thus some form of regulation or a regime to combat concerns over security (Altmann & Sauer 2017: 132; Altmann 2013: 137; 2009: 69). Moreover, authors often draw comparisons between AWS and nuclear weapons, and arms competition during the Cold War (Altmann & Sauer 2017: 118; Scharre 2017: 19–20; Geist 2016: 318–319). Recently, authors have addressed the impact of AWS use on regional security such as in the Asia-Pacific (see Leys 2018; Koh 2017). Other studies have focused on the normative change brought by AWS, relying on critical theories, discourse or cultural analysis (see Karppi et al. 2018; Beier 2017; Ekelhof 2017; Haas et al. 2017).

However, as Bode and Huelss (2018) suggest "[...] the autonomy of weapon systems is accelerating largely outside of public and academic attention in the discipline of International Relations [...]" (Bode and Huelss 2018: 1–2). In fact, no articles on AWS explicitly draw from theories of International Relations with the exception above. Thus, this lack of theoretical clarity combined with overlapping arguments across different approaches constitutes a severe obstacle in systematic mapping of the research on AWS. Furthermore, it seems clear that AWS present a polarizing issue as demonstrated by the "roboutopia" vs "robocalypse" divide. A call for action, typically some form of regulation, is usually justified based on perceived severity of issues in question. It should be noted that the goal of this thesis is neither to argue for or against AWS nor assess which approach to AWS is the best. The goal is to assess which one of the potential solutions, or outcomes, of the current debate is the most likely. In this context, the thesis bridges the gap in literature by creating a theoretical framework which explicitly draws from IR and incorporates contributions of other approaches, while introducing the role of state preferences and power.

3.2. Regulatory Approaches to AWS: From Minor Regulation to Total Prohibition

Before elaborating on theoretical assumptions, it is important to provide an overview of the potential policy solutions to AWS found in the literature. These can be divided into pro-ban and pro-regulation approaches in general (see Dunkelberg 2016; Geist 2016). Some authors prefer distinguishing between soft and hard law, and soft and hard governance (see Crootof 2016a; Crootof 2015; Marchant et al. 2011). However, with regard to the sheer number of particular solutions proposed so far, this general distinction does not hold a great analytical value. To provide a clearer understanding, the section will distinguish between five approaches: (1) no regulation, (2) minor regulation, (3) hybrid regulation, (4) partial prohibition, and (5) total prohibition. The review organized like this will allow keeping the general distinction between ban and regulation approaches, while establishing foundations for operationalization of AWS policy dimension, which will serve to capture states' preferences. The aim is to capture the variation in the degree of perceived severity of the challenges posed by AWS, or willingness of states to commit themselves to restrictive arrangements.[21]

Total prohibition. A substantial part of discussions focuses on the prospects of a "full ban" and a "pre-emptive ban" (see Dunkelberg 2016; Goose 2015; Wallach 2017b). Such ban would make AWS illegal under international law, regardless of the circumstances of their use, such as offensive or defensive purpose, and simple or complex environments (Marchant et al. 2011: 300–305). The proponents of the approach include the Human Rights Watch (HRW) and multiple other NGOs under

[21] To be more precise, it is the variation in the degree of "obligation" as part of the legalization concept described by Abbott et al. (2000) (see Abbott et al. 2000).

the banner of the Campaign to Stop Killer Robots (CSKR) (Geiss 2015: 23). Some states have also expressed support to prohibition of AWS (CSKR 2018b). However, authors often doubt the likelihood of the success of such approach because most technologically advanced states explicitly refuse total prohibition. In addition, a treaty regime without states with incentives to obtain AWS would likely fail to achieve the effect (Dunkelberg 2016: 8–9). Arguments in favor of the "full ban" approach were formulated in the 2012 HRW report titled "Losing Humanity". As the title suggests, the report argues that the use of AWS will be fundamentally incompatible with principles of IHL. The report further focuses on ethical challenges posed by AWS (HRW 2012: 1–4). The approach is seen as appropriate in part because it would be clear and easy to enforce (Crootof 2016a: 1896–1897).

Partial prohibition. Besides the "full ban" approach, authors suggest instead to ban specific uses or types of AWS (Dunkelberg 2016: 22–24; Marchant et al. 2011: 299–305). With regard to specific legal solutions, authors suggest adopting a treaty limiting the use of AWS to non-human targets, defensive purposes, conflicts among states, or, perhaps, large battlefields (Dunkelberg 2016: 22–23; Cass 2015: 1046–1047). In addition, prohibition may be imposed on certain payloads such as nuclear weapons (Marchant et al. 2011: 298). However, prohibition can also apply to specific stages of acquisition, R&D, testing, deployment, transfer, proliferation, and so forth (Gubrud & Altmann 2013: 3; Marchant et al. 2011: 299). Similarly, the involvement of technologically advanced states would be necessary for legally-binding treaty to be effective but the solution is less restrictive, and potentially more acceptable (Dunkelberg 2016: 8–9). The standard of meaningful human control (MHC) typically comes up in this context as a principle to prohibit weapons using force without human oversight.[22] The concept was invented by NGO Article 36 as a measure to preclude accountability gaps (Article 36 2014: 1). However, while the notion has gained wide support among many states, proposed forms of MHC regulation differ immensely (Geiss 2015: 24). Some suggest adopting MHC as a basic principle of IHL by amending Additional Protocol I or the preamble of the CCW, others understand MHC as more of an informal standard or guidance (Rosert 2015: 1; Neslage 2016: 169–171; Geiss 2015: 24).

[22] Sometimes the terms "sufficient human control" or "appropriate human control" are used as well (Article 36 2014: 1).

Hybrid regulation. While prohibition often relies on hard law, and therefore legally binding treaties, hybrid regulation here refers to a mix of soft and hard law regulation. This includes forms of regulation where states retain significant powers to interpret and implement the rules, while acknowledging both risks and benefits of AWS (Altmann & Sauer 2017: 134; Jenks 2016: 54; Marchant et al. 2011: 314). Internationally agreed moratorium and framework conventions present examples of such regulation (see Crootof 2016a; Marchant et al. 2011). Additionally, it is possible to distinguish between a “comprehensive” and “limited” moratorium, the former constituting more prohibitive form (see Jenks 2016). It was suggested, for example, by Christof Heyns (2013) at the HRC. States could impose moratoria on testing, production, assembly, transfer, acquisition, or deployment and use of AWS (Heyns 2013: 21). Though the solution would be less restrictive than a ban, it is questionable if countries would be willing to implement a “blanket” moratorium (Dunkelberg 2016: 21; Jenks 2016: 59–62). Nevertheless, no state has publically endorsed the approach (Geiss 2015: 24). By developing a framework convention, this would allow countries to adopt protocols on particular issues on a voluntary basis without strong commitments (Crootof 2016a: 1899; Marchant et al. 2011: 313).

Minor regulation. Aside from prohibitive instruments, a wide variety of soft, non-binding legal instruments has been suggested (see Crootof 2016a; Marchant et al. 2011). This includes resolutions or declarations, international and transnational dialogue, best practices, information-sharing, confidence-building measures so as to improve communication, transparency, and verification, codes of conduct (see Crootof 2016a; Marchant et al. 2011). An important advantage of these solutions is that they can be adopted with relative ease compared to some of those mentioned above, while allowing for greater flexibility and adaptability to changes in AWS R&D (Dunkelberg 2016: 9; Marchant et al. 2011: 306). Even institutions such as HRW have suggested professional codes of conduct developed by roboticists and others involved in AWS R&D as a starting point (HRW 2012: 5). Such codes are usually developed by technical experts, private companies, and others, as a way to describe professional ethics and best practices in a specific issue-area (Crootof 2016a: 118; Marchant et al. 2011: 307). For this reason, they are potentially most likely related to technical approaches to AWS. Moreover, minor regulation would likely be preferred by those optimistic about prospects of AWS, or those weighing benefits over costs.

No regulation. Lastly, even though there are currently no AWS-specific laws and treaties, this does not mean that AWS would be completely unregulated. Many bodies of law would apply even in the absence of such regulation. For example: IHL and IHRL, Law of the Sea, Space Law, customary law, Article 36 weapons reviews, and international legal regimes such as the Missile Technology Control Regime (Crootof 2016a: 98; Marchant et al. 2011: 289). Furthermore, there is a variety of universal constraints to misuse of weapons found in military doctrines and codes of ethics (Marchant et al. 2011: 290–293). Eventually, precedents might emerge which could develop into customary rules, thereby compensating for the lack of specific regulation (Crootof 2015: 98). This opinion is held by a relatively small group of legal and military experts who perceive compliance mostly as "a matter of technological development" (see Neslage 2016; Cass 2015; Anderson & Waxmann 2013; Schmitt & Thurnher 2013). Lastly, the same sense of technological optimism is reflected in contributions by some roboticists, including Arkin and his "ethical governor" software (see Arkin 2017; 2010; Arkin et al. 2009). Proponents of this approach interpret efforts to ban AWS as irresponsible due to the potential benefit it can bring (Geiss 2015: 25).

4. Theory

This chapter aims to establish a theoretical framework for analysis in order to answer the research questions formulated in the introduction. First, this chapter provides aa overview of defensive realism—one of the key paradigms of IR and an intellectual foundation for arms control. The section discusses contributions of the most prominent defensive realists, including Kenneth Waltz, Robert Jervis, and Charles Glaser (see Glaser 1995; Waltz 1979; Jervis 1978). In particular, it focuses on concepts of international anarchy, power, security dilemma, arms-racing and its drivers, offense-defense balance, as well as conditions for cooperation and critique of neorealism. Second, the following section provides overviews of traditional and preventive arms control approaches, focusing on the issue of bargaining over arms control agreements, and their legalized forms. The section also provides empirical-conceptual syntheses, which connect the empirical background with the theory. As for the last section, third, it formulates hypotheses drawn from the framework, and elaborates on dependent and independent variables, including state preferences of AWS regulation, and institutional activity in AWS discussions.[23]

4.1. Defensive Realism: Security-Seekers in Anarchical International System

Beside neoliberalism and constructivism, neorealism, or structural realism, constitutes one of the dominant paradigms of IR theory. In essence, it is a theory of power politics, which is generally pessimistic about prospects of cooperation and elimination of war from the repertoire of international affairs (Walt 1998: 31). The basis of the neorealist theory was laid down by Kenneth Waltz in his book "Theory of International Politics" (see Waltz 1979). In comparison with the classical realist

[23] There are two reasons for incorporating additional approaches besides defensive realism. First, while defensive realists provides the broader underlying assumptions, the analysis requires more in-depth understanding of issues pertaining to arms control—a subject which lies at the heart of the negotiations, whether formal or informal, of AWS regulation. Second, the finding what specific form of regulation is likely to prevail requires deeper understanding of the regulatory dimension, which is explored in contributions of institutionalist literature. Incorporating the institutionalist approach is possible because the subject of legalized agreement merely reflects the empirics.

theory, which focused on the human nature, Waltz (1979) focused on the effects of international system (Walt 1998: 31). Opinions on key assumptions of neorealism may vary but most neorealists agree on several of them. First, international system is populated by functionally similar units—states—and its main ordering principle is anarchic, meaning that there is no central governing authority. Second, states are rational and unitary actors, acting in self-interest. Third, states have some military capability, which can be used to coerce or destroy. Fourth, states cannot be certain of others' intentions. Fifth, states' most basic motive is survival (Lobell 2010: 6651; Glaser 1995: 54; Mearsheimer 1995: 10–12).[24]

Due to the aforementioned effects, states tend to fear each other, each seeks to ensure its survival through self-help, and maximize its relative power compared to others (Mearsheimer 1995: 11–12). Self-help since there is no authority capable of protecting states, and the resulting competition among states is catalyzed by the uncertainty about others' intention. Consequently, conflict rather than cooperation characterizes interaction among states (Glaser 1995: 55–56). Neorealism suggests that states prefer competitive policies such as arms build-up to alliances and arms control to meet their security demands, particularly, due to concerns over cheating and relative gains (Glaser 1995: 56; Mearsheimer 1995: 12). As Grieco (1988) puts it, "[...] states are positional in character" (Grieco 1988: 600). Therefore, when joint gains favor state's partners, distributional issues may rule out cooperation (Grieco 1988: 600). Neorealism therefore predicts cooperation to be rare, limited to areas of secondary importance (Glaser 1995: 57). Nevertheless, not all neorealists are so skeptical about the prospects of cooperation (see Jervis 1978; Glaser 1995).

As an extension of Waltz's structural realism, neorealism can be subdivided into offensive and defensive realism. The crucial distinction between the two rests in the implications of international anarchy (Lobell 2010: 6651). Offensive realists argue that the international anarchy incentivizes states to maximize power as only the most powerful states in the system can be sure of their survival. Consequently, offensive realism understands security to be scarce (see Mearsheimer 2001; Gilpin 1981). In contrast, security is said to be plentiful from the perspective of defensive

[24] According to Kenneth Waltz: "In anarchy, security is the highest end. Only if survival is assured can states safely seek other goals as tranquility, profit, and power" (Waltz 1979: 126).

realism. International anarchy incentivizes states to pursue modest policies rather than conquest. For defensive realists, states are essentially security-seekers rather than power-maximizer (see Glaser 1995; Waltz 1979; Jervis 1978). This is because aggressive policies provoke counterbalancing and security dilemma (Lobell 2010: 6652; Taliaferro 2000: 129). Thus, states mainly seek to obtain only the necessary means to survival rather than superiority (Walt 1998: 31). Consequently, defensive realists are more optimistic about the prospects of cooperation (see Glaser 1995).

One of the central concepts to all factions of realism is the concept of power. Perhaps unsurprisingly, it is also one of the most contested concepts in IR theories (Nye 2011: 10; Rothman 2011: 49). For example, Dahl (1961) describes power as a means of getting what we want from others contrary to their initial preference (see Dahl 1961). From the neorealist perspective, power comprises of components such as "[...] size of population and territory, resource endowment, economic capability, military strength, political stability and competence" (Waltz 1979: 131). Defensive realists such as Glaser (1995) focus on military capabilities, and ability to perform military missions (Glaser 1995: 58). Traditional realist "hard" power comprises of material factors, particularly economic and military resources, and can be broadly defined as power of coercion (Nye 2011: 19; Rothman 2011: 52; Wilson 2008: 114; Walt 1998: 40). However, pragmatic realists often consider other power resources as well, including ideas and persuasion—components of "soft" power that relies on means such as framing the agenda or positive attraction (Nye 2011: 18).[25]

One of the central concepts of defensive realism is called "security dilemma" (see Herz 1950; Jervis 1978; Glaser 1997; Tang 2009). This phenomenon was first descried by John Herz, who observed that under conditions of anarchy, whenever a state seeks to increase its security, it renders the others more insecure, compelling them to prepare for the worst (Herz 1950: 157). According to Robert Jervis (1978), the dilemma ensues if "[...] many of the means by which a state tries to increase its security decrease the security of others" (Jervis 1978: 186), or, similarly, if "[...] an increase in one state's security decreases security of others [...]" (Jervis 1978: 186).

[25] The concept of soft power was developed by Joseph Nye in 1990. Soft power differs from hard power in that it does not rely on coercive means but rather on persuasion and attraction. It includes intangible factors such as institutions, ideas and values rather than material factors. Nye also developed the concept of "smart" power, which combines soft and hard power (Nye 2011: 19–20).

This logic provides the basis for the “spiral” model, which explains how states with essentially compatible goals can end up in a conflict (Jervis 1999: 49). To illustrate this dynamic, Jervis (1978), employs a simple 2x2 prisoner’s dilemma game, which reveals how the phenomena operates, and what steps might be taken to avoid such mutually deficient outcomes such as arms races and wars (see Jervis 1978).[26] The figure below depicts the 2x2 game (see Figure 8).

Figure 8: Prisoner’s Dilemma

	Player B	
Player A	C Cooperation	D Defection
C Cooperation	R=3 Reward for mutual cooperation	S=0 Sucker's payoff
D Defection	T=5 Temptation to defect	P=1 Punishment for mutual defection

Fig. 1. The Prisoner's Dilemma game. The payoff to player A is shown with illustrative numerical values. The game is defined by $T > R > P > S$ and $R > (S + T)/2$.

Source: Axelrod and Hamilton (1981: 1391)

In prisoner’s dilemma, if the game is played only once, the rational choice is to defect, despite the potential gains from mutual cooperation (Axelrod & Hamilton 1981: 1391; Jervis 1978: 171). However, if the game is played repeatedly, then this no longer holds, and mutual cooperation is possible under certain conditions. What drives the prisoner’s dilemma is ultimately the fear of being exploited (CD). States, which are relatively more powerful, and would suffer lower costs of being cheated, have, however, a margin of time and error, which dampens their security dilemma. When the cost of (CD) is low, relatively low level of arms suffices and states adopt non-threatening policies. Nevertheless, how players perceive (CD) may be affected

[26] For an extensive definition of security dilemma, see Shiping Tang (2009). In addition, the author observes that although various realist contributions suggest many different aspects of security dilemma, three aspects in particular form the basis of a genuine security dilemma. These include, first, anarchy which leads to uncertainty and fear among states, second, a lack of malign intentions, and third, accumulation of power (Tang 2009: 593–595).

by subjective security demands, including feeling of vulnerability, and perceptions of others as a menace. Probability of mutual cooperation (CC) can be improved, for example, by increasing the gains from (CC), decreasing the cost of (CD), decreasing the gains from (DC), or increasing the cost of (DD). Inspection devices can be used as a means of ameliorating the dilemma (Jervis 1978: 171–175, 181).[27]

When conditions for cooperation are lacking, repeated mutual defection can fuel arms competition and arms races (see Bolks & Stoll 2000; Kydd 2000; Morrow 1989; Downs et al. 1985; McCubbins 1983; Richardson 1960). An arms race can be described as a conflictual relation among states competing over the strength of the armed forces both in quantitative and qualitative terms, meaning the size and level of sophistication of their arsenal (see Huntington 1958). While the spiral model or repeated prisoner's dilemma suggest adopting cooperative policies to reduce risks of arms-racing turning into wars, the deterrence model suggests arms competition as a way to avoid war itself (Glaser 2000: 251; Kydd 2000: 228–229). Spiral model views arms-racing as an action-reaction process, where build-up of arms generates insecurity and misperceptions that feed into a vicious circle (Schofield 2000: 760). Literature on arms races suggests both international and domestic causes of arms-racing. These include: existence of external threats, interest of the R&D or military-industrial complex, and others (Glaser 2000: 251, 257; Buzan & Herring 1998: 83). A state may also seek to build arms because it seeks expansion—it is greedy (Glaser 2000: 254).[28]

Many other explanations why states seek to procure arms can also be found in the arms trade literature. Pearson (1989) suggests variables of national, regime, military, economic, security, and alignment-pattern indicators (see Pearson 1989). Some constitute important economic enablers such as GDP per capita, and military expenditure (Martínez-Zarzoso & Johannsen 2017: 16; Pearson 1989: 154–155), or

[27] Glaser suggests two additional variables which affect the magnitude of the security dilemma. These are: the adversary's greed, and adversary's unit-level knowledge. Knowing whether a state is facing greedy adversary or a security-seeker can be essential for choosing between cooperative and competitive policies (Glaser 1997: 174).

[28] Lakoff and Bruvold (1990) observe that, from a historical perspective, qualitative arms races tend to be difficult to control, which feeds into the widely held belief in "deterministic forces of the scientific progress" (Lakoff & Bruvold 1990: 389). Some authors have questioned whether qualitative arms races are driven by conscious decision-making, or whether the technological progress itself drives decision-making (Hamlett 1990: 462–463; MacKenzie 1989: 162–163; Thee 1986: 127).

political predispositions such as democracies and autocratic regime (Fuhrmann & Horowitz 2017a: 410; Pearson 1989: 154), and militarization of society (Martínez-Zarzoso 2017: 16). These are some of the factors that may incentivize countries to build arms of a nature detrimental to the security dilemma.

Figure 9: Offense-Defense Balance and the Security Dilemma

	OFFENSE HAS THE ADVANTAGE	DEFENSE HAS THE ADVANTAGE
OFFENSIVE POSTURE NOT DISTINGUISHABLE FROM DEFENSIVE ONE	1 Doubly dangerous	2 Security dilemma, but security requirements may be compatible.
OFFENSIVE POSTURE DISTINGUISHABLE FROM DEFENSIVE ONE	3 No security dilemma, but aggression possible. Status-quo states can follow different policy than aggressors. Warning given.	4 Doubly stable

Author: Jervis (1978: 211)

How the security dilemma operates depends on the offense-defense balance (see Glaser & Kaufmann 1998; Van Evera 1998; Levy 1984; Jervis 1978). When one country increases its security by pursuing only distinctively defensive policies and weapons, the dilemma need not operate so severely. The issue arises when offense has the advantage—that is, attacking is effectively easier than defending—and when defensive and offensive postures are indistinguishable. If these conditions become true, states are likely to engage in arms races, and fear of reciprocal surprise attack will lead to strategic instability, incentivizing preemption. Distinguishing between status-quo states and aggressors then becomes very difficult, and cooperation very unlikely. Conversely when defense has the advantage and defensive posture can be differentiated from offensive, states can attain security with relatively low levels of arms and with compatible policies (Jervis 1978: 186–187). Jervis (1978) identifies four situations arising from the offense-defense balance (see Figure 9). Dominance of either is caused by factors such as military technology and doctrine, geography, social structure, diplomacy, and others (Van Evera 1998: 6).

Glaser and Kaufmann (1998) identify six major areas of technology relevant to the offense-defense balance, including mobility, fire-power, protection, logistics, communication, and detection (Glaser & Kaufmann 1998: 62). Weapons depending on surprise attacks are usually identified as offensive (Jervis 1978: 206). However, assessment of any weapon in offense-defense terms is often very complicated. For Shiping Tang (2010), the differentiation is meaningless because most weapons are inherently dual-use. He suggests that "objective" offense-defense balance is a hoax, although perceptions of offensive-defensive advantage might still play a role (Tang 2010: 222, 240). From an empirical standpoint the assessment of AWS presents an issue for various reasons. First, autonomy should be treated as a characteristic of a weapon rather than anything else, and as such could improve upon all of the above mentioned areas of technology, supporting both offense and defense (see US Army 2017; DoD 2013). Second, systems that enjoy higher degrees of autonomy have so far been deployed only for defensive purpose, thus offensive capability can only be assessed in terms of hypothetical scenarios (Altmann & Sauer 2017: 118; Boulanin & Verbruggen 2017: 115). Third, the offensive-defensive advantage comes down to software, which can be quickly updated, hacked, or manipulated (Altmann & Sauer 2017: 120; UNIDIR 2017a: 10–12). Paradoxically, AWS can be purely offensive and defensive at the same time.[29]

As noted earlier, the nature of the security dilemma and the offense-defense balance is relevant to the repertoire of policies available to states (see Glaser 1997; 1995; Jervis 1978). When offense and defense are indistinguishable, states have no guidance on whether to pursue cooperative or competitive policies. However, they can signal benign intentions by engaging in arms control negotiations, by adopting unilateral defense, or unilateral restraint. Glaser (1995) argues that arms control is largely unnecessary when defense has a large advantage. Conversely, when offense

[29] Various authors suggest that AWS provide significant first-strike advantage due to advanced collaborative behavior such as "swarming" and ability to conduct saturation attacks to overwhelm the opponent. Whether offensive or defensive swarms would have an advantage is questionable. Nonetheless, no offensive swarms have yet been deployed (see Altmann & Sauer 2017; Lachow 2017).

has an advantage, arms control is more useful. As the advantage of offense further increases, the case for arms control becomes less clear because incentives to cheat will be high. When offense and defense is distinguishable, the need for precise and effective verification decreases, and so do incentives to cheat. In the opposite case, arms control is less useful. Furthermore, costs of arms control are typically higher for "greedy" states compared to genuine security-seekers when the offense has an advantage. Therefore, when countries agree on arms control under conditions of indistinguishability, this can communicate information about motives—greedier states are less likely to adopt such agreement (Glaser 1995: 65–69).

Neorealists generally agree that international institutions do "matter" in the context of cooperation, whether related to arms control or other areas, although to a lower extent than neoliberalist and constructivist paradigms would suggest (see Glaser 2003; Mearsheimer 1995; Jervis 1999). International institutions, according to Mearsheimer (1995), can be defined as "[...] a set of rules that stipulate the ways in which states should cooperate and compete with each other [...]" (Mearsheimer 1995: 8).[30] Jervis (1999) provides a very accurate neorealist take on institutions, which are essentially "[...] the product of the same factors—states' interests and the constraints imposed" (Jervis 1999: 54). In contrast to other paradigms, institutions are understood merely as tools of statecraft, a means of states to achieve their own interests, rather than as autonomous, independent actors. Institutions are believed to reflect states' interests, and, particularly, interests of great powers (Glaser 2003: 409; Mearheimer 1995: 13; Jervis 1999: 43). According to proponents of defensive realism, institutions can serve to lower transaction costs or improve transparency and thus ameliorate the fear of exploitation but do not change states' motives from self-interest to altruism (Jervis 1999: 52; Glaser 1995: 84).[31]

Structural realism is often criticized for neglecting if not outright dismissing the role of domestic-level variables and processes, thereby treating states as "black

---

[30] Other commonly used definitions include, for example, "[...] enduring sets of rules, norms, and decision-making procedures that shape the expectations, interests, and behavior of acts [...]" (Goldstein et al. 2000: 387).

[31] In addition, a fairly common perspective consistent with the neorealist definition is to study institutions as forums for interstate bargaining (Glaser 2003: 409; Fearon 1998: 298; Morrow 1994a: 408–411).

boxes" (Sterling-Folker 1997: 1, 16; Zakaria 1992: 180). By the same token, critics highlight the hypocrisy of looking to the domestic level only once unable to explain state behavior with structural causes (Rathbun 2008: 296). However, this criticism is somewhat unwarranted. Many realists, such as Kenneth Waltz and Glenn Snyder, acknowledge the importance of domestic variables (see Snyder 1991; Waltz 1979). Waltz (1979) suggests that "[e]ach state arrives at policies and decides on actions according to its own internal process [...]" (Waltz 1979: 65). Some realists further suggest that systemic and domestic-level theorizing is compatible and a necessary extension of realism (see Sterling-Folker 1997). Ultimately, it is the domestic actor who makes the decision within the domestic context, including political pressures, ideological preferences, or access to information, and by which it reacts to external conditions (Sterling-Folker 1997: 16–17; Zakaria 1992: 180).[32]

Beyond the domestic-level influences debate, neorealism has been criticized for many other reasons (see Taliaferro 2000; Legro & Moravsik 1999; Wendt 1999; Ashley 1984). Legro and Moravcsik (1999) argue that neorealism lost explanatory value as a result of focusing excessively on anomalies, and recasting the theory in a way, which better explains such anomalies by "diluting", rather than deepening, its fundamental assumptions (Legro & Moravcsik 1999: 6, 44). The authors thus label various neorealist factions as "minimal realism" because they only retain the very fundamental assumptions of neorealism—international anarchy, and rational state-actors (Legro & Moravcsik 1999: 6). Furthermore, they criticize Glaser for shifting the emphasis from the neorealist causal reliance on the distribution of capabilities to the international institutional environment, which is "overtly functional". Legro and Moravcsik essentially suggest that Glaser's "contingent" realism, as he puts it, depends on assumptions of rational institutionalism and is not distinctively realist (Legro & Moravcsik 1999: 43–45). This thesis acknowledges that there are various similarities between the approaches but it ultimately does not seek to address this criticism.

[32] The impact of domestic processes explains why the choices in response to environment are not identical. As Waltz (1979) puts it, structural realism predicts state behavior under same structural conditions to be "similar" not "identical" (Waltz 1979: 122).

4.2. Arms Control: Reaching Agreements under the Security Dilemma

As discussed in the previous section, states can ameliorate security dilemma and avoid costly arms races by adopting arms control agreements (see Jervis 1978; Glaser 1995). Generally, arms control can be defined as "[...] any agreement among states to regulate some aspect of their military capability [...]" (Larsen 2002: 1). In addition, the term can refer to "[...] a series of alternative approaches to achieving international security through military strategies" (Selim 2013: 7). Arms control as a theory emerged during the Cold War as a result of the US-Soviet arms race (Selim 2013: 7). Some of the key contributions included Schelling and Halperin's "Strategy and Arms Control", and Hedley Bull's "The Control of the Arms Race" (see Schelling & Halpering 1985; Bull 1961). Shelling and Halperin (1985) identified objectives of arms control as reducing the likelihood of war, reducing the political and economic costs of preparing for war, or minimizing the scope of war once it occurs (Schelling & Halperin 1985: 3). Bull (1976) also mentions containing proliferation, stabilizing relations between great powers, or combating militarization of society (Bull 1976: 5–6).[33]

Furthermore, the "classical" or "traditional" arms control can be understood as a means of enhancing national security (Bull 1983: 21). As such, it does not aim at disarming states completely, depriving them of the means to defend themselves. Disarmament should be seen as conceptually different from arms control. The first seeks elimination of certain classes of weapon, whereas the latter seeks limitation, on acquisition, production, deployment, or use of weapons. Additionally, the latter does not rule out the possibility of allowing states to procure more arms. National security takes precedence over the reduction of arms (Selim 2013: 7; Larsen 2002: 3). Therefore, arms control is largely compatible with defensive realist thinking of states as security seekers. If the agreements bring states equitable mutual benefits, adequate monitoring is possible, and commitments are credible, they can promote and maintain stable balance of power among signatories (Larsen 2002: 5; Lakoff &

[33] Moreover, some authors stress the need for re-labeling of the term partly as an attempt to broaden the scope so it encompasses more neatly the activities associated with control of arms that have a direct effect on society, R&D or dual-use technologies (Cooper & Mutimer 2011: 10).

Bruvold 1990: 390). However, conclusion of such agreement is usually lengthy and difficult process.[34]

The main issue in concluding any arms control agreement can be illustrated on the previously described 2x2 prisoner's dilemma game. Despite the potential of mutual gains from cooperation, each state has an incentive to defect due to fears of being cheated by others (see Oye 1985; Axelrod & Hamilton 1981). The problem is especially severe in case of arms control since constraints on arms could leave one vulnerable to attack (Hansel et al. 2018: 5–6; Abbott & Snidal 2000: 435; Bernauer & Ruloff 1999: 10). As a result, ensuring that nobody cheats on their commitments requires effective monitoring and enforcement mechanisms (Hansel et al. 2018: 5; Albin & Druckman 2014: 434). Nevertheless, the 2x2 prisoner's dilemma game can be somewhat misleading. In practice, there are many possible ways in which arms control agreements can be struck, each with different distributional consequences as regards costs and benefits. Thus, Fearon (1998) observes that parties first have to solve the distributional problem, agreeing on which of the potential outcomes of cooperation would be enforced (Fearon 1998: 273–274).[35]

When parties prefer different kinds of solutions to the dilemma, this creates a bargaining problem (Koremenos et al. 2001: 773). There might be many possible alternatives to no agreement but countries' perception of costs and benefits as well as the order in which they prefer these alternatives differ (Bernauer & Ruloff 1999: 10; Fearon 1998: 274; Morrow 1994a: 388). Thus, states engage in the process of bargaining, exchanging offers and counteroffers, while seeking concessions from others to reach their most preferred outcome (Fearon 1998: 274; Steinberg 1985: 262). In the context of arms control, asymmetries in preferences usually arise from differences in quantity or quality of military hardware, scheduled procurement, or budgetary and administrative constraints to implement the agreement. In addition, some authors argue that countries seeking economic liberalization are more likely to engage in negotiation, and states facing external threats are the least likely to do

[34] Various attempts at banning certain weapons or methods of warfare have failed in the past. Examples include aerial bombardment, nuclear weapons, or submarines (Crootof 2016a: 1904–1909).

[35] It should be emphasized that, in line with realist argumentation, it is the concern over "relative" distribution of costs and benefits from the agreement, which make or break the agreement (Schofield 2000: 756).

so (Bernauer & Ruloff 1999: 11–16). Differences in technological level exacerbate the problem because banning certain technologies might disproportionally benefit those lagging behind, while costing leaders in R&D (Hansel et al. 2018: 6).

One aspect in which classical arms control differs markedly from defensive realism is the concept of "bargaining power". While realists would suggest that the agreement reflects the interests of those with the largest share of power as regards military capability, Schelling (1956) further emphasizes the role of psychology and context (Lebow 1996: 555; Schelling 1956: 282). Lebow (1996) summarizes three other influences, besides "hard" power, highlighted by Schelling. First, the context, which determines the stakes, range of possible outcomes, issue salience, or ability of states to commit themselves to particular outcomes. Second, skill, which affects the ability of states to improve their position, for example, by minimizing the cost of backing down, or discrediting others' commitment. And third, the willingness to suffer, which makes it more difficult for others to achieve concession from oneself (Lebow 1996: 556–558). Paradoxically, appearing weak, or unable to change one's own position, for example, because of domestic pressures, can increase bargaining power (Schelling 1956: 288).

One strand of arms control focuses specifically on prevention of qualitative arms races (see Neuneck & Mutz 2000; Altmann 2008; Mutschler 2013). According to Hansel et al. (2018), "preventive arms" control seeks to avoid arms races during the early stages of R&D, and preventing it from being deployed (Hansel et al. 2018: 1). Altmann et al. (1998) note that preventive arms control is "[...] qualitative arms control applied to the (near or far) future" (Altmann et al. 1998: 256). Accordingly, proponents of the approach tend to focus on the effect of military R&D on stability and security dilemma. Preventive arms control calls for anticipatory assessment of technological developments and adoption of preventive measures in order to avoid deployment and use of weapons with destabilizing potential (Hansel et al. 2018: 1; Altmann et al. 1998: 256). Justification for such preventive measures comes, again, mostly from lessons of the US-Soviet nuclear competition, particularly with regard to qualitative improvements which sought a decisive advantage—these include, for example, MIRVed missiles or anti-ballistic missile defense systems (Altmann 2001: 244; Altmann et al. 1998: 255).[36]

Proponents of preventive arms control typically argue that traditional arms control has not paid enough attention to those qualitative improvements, focusing instead on adoption of quantitative limits. They also argue that their approach goes beyond qualitative arms control in that it seeks to address the dangers of spread of the technology (Altmann et al. 1998: 256). This process is characterized by several hurdles specific to such efforts. Crucially, most technologies during the R&D phase are inherently dual-use. This poses a significant challenge to verification, and total prohibition is almost never possible without repercussions to civilian innovation and non-military use (Hansel et al. 2018: 6; Altmann et al. 1998: 262–263). Despite these issues, several agreements exist which include elements of preventive arms control. These are, for example, the Biological and Chemical Conventions, the now-defunct ABM Treaty of 1972, nuclear testing treaties, and the Protocol on Blinding Laser Weapons of 1995 (Altmann 2004: 4–5).

Preventive arms control offers a succinct framework for assessing whether the technology should be preventively prohibited (see Altmann 2009; 2004; 2001; Neuneck & Mölling 2001). The Framework encompasses both “classical” criteria of arms control such as prevention of war, damage avoidance, and costs reduction, as well as criteria of sustainable development or human security (Neuneck & Mölling 2001). When criteria are satisfied, and the technology thus proves to be dangerous, preventive arms control suggests states to negotiate an agreement (Altmann 2009: 73; 2004: 71; Altmann 2001: 244). Full list of criteria identified by Altmann (2009) can be found in the appendix (see Appendix 1). While the first and second group of criteria reflects “classical” arms control approach, the third group further includes societal and political challenges. Drawing from the framework, the author provides a modified list of issues in AWS regulation, which serve as an empirical-theoretical synthesis (see Table 3). Thus modified criteria will allow a tailor-made estimation of state preferences.

[36] Examples of application of preventive arms control approach include studies by Hansel et al. (2018) on cyber weapons, Altmann and Sauer (2017) on AWS, Mutschler (2013) on space weapons, Altmann (2013, 2009) on UVs and partly AWS, Altmann (2004) on nano-technology, or Altmann (2001) on non-lethal weapons technologies (see Hansel et al. 2018; Altmann & Sauer 2017; Mutschler 2013; Altmann 2013; 2009; 2004; 2001).

Table 3: Empirical-Conceptual Synthesis of Issues in AWS Regulation[37]

Issue-area	Main sources of influence	Criteria of preventive arms control	Examples of specific issues from empirical background
International law	External environment and domestic processes	*Adherence to and further development of effective arms control, disarmament and international law*	Potential to improve compliance with international law and forms of regulation vs. likelihood of indiscriminate attacks, disproportionate attacks, accountability gaps, and feasibility of Article 36 reviews
International security	External environment	*Maintain and improve stability*	Potential military benefits, including force substitution, costs reduction, dull, dirty, and dangerous missions, and casualty reduction vs. impact on arms-racing, proliferation, strategic stability, asymmetrical warfare, and "the strategic robot problem"
Societal security	Domestic processes	*Protect humans, environment and society*	Potential to improve ethical conduct on the battlefield vs. violation of human dignity, public backlash, exploitation of "risk-free" warfare, devaluation of military service, and moral disengagement
Operational safety	Technology	Absent as an individual category; implicitly included in above mentioned criteria	Potential to improve performance, including speed, accuracy, precision, or secure data links vs. risks stemming from complexity, predictability, reliability, vulnerability, and trust in human-machine relationship

Source: Author (2018)

[37] While states, according to defensive realism, give clear priority to security in terms of sensitivity to external threats, Glaser (2003) argues that, once states have achieved satisfactory level of security, pursuit of other objectives, for example humanitarian and ideological interests—included here under labels international law and societal security—does not violate structural realist assumptions. Similar argument applies to tradeoffs between domestic objectives and security as describe above (Glaser 2003: 412–413). Technology as a separate source of influence is considered due to its significance in states' assessment of the offense-defense balance but also because technological progress itself can sometimes drive states' ambitions (Hamlett 1990: 462–463; MacKenzie 1989: 162–163; Thee 1986: 127).

The degree of commitment required by a treaty can be measured by looking at the levels of legalization of particular solutions to AWS regulation. This concept was introduced by Goldstein et al. (2000) and elaborated in several institutionalist contributions (see Abbott et al. 2000; Abbott & Snidal 2000; Goldstein et al. 2000). Before elaborating the concept itself, however, it is important to explain how it fits into theoretical assumptions of defensive realism. First, the concept is empiricist in origin. It simply reflects existing types of legalized agreements (Abbott et al. 2000: 402). Thus, application of the concept on its own does not depend on assumptions of institutionalist theory. That being said, second, caution must be exercised when drawing conclusions about the nature of state preferences towards such degrees of legalization. These should be seen as pre-existing preferences that result primarily from structural conditions or sates' calculation of costs and benefits, rather than as a result of a process of legalization brought about by independent non-state actors, such as international institutions.[38]

According to Goldstein et al. (2000), legalization should be understood as "a particular form of institutionalization" (Goldstein et al. 2000: 386). Different forms of legalization stem from variation in three criteria, including obligation, precision, and delegation. The first refers to the extent to which states are bound by rules and commitments set out in the agreement, the second refers to the extent of clarity or unambiguity with which such rules are defined, while the third refers to the extent to which third parties exercise authority to implement the agreement (Abbott et al. 2000: 401; Goldstein et al. 2000: 387). Importantly, each of these above mentioned dimensions constitutes a continuum rather than dichotomous variable. As a result, the overall degree of legalization comprises a multidimensional continuum, which encompasses several "ideal types", including "hard law", "soft law" and absence of

[38] With respect to the obligation aspect of legalization, in particular, Abbot et al. (2000) claim that obligation, as in "legal" obligation, differs from obligations resulting from coercion (Abbot et al. 2000: 408–409), while realists typically argue that legal obligations embodied by a legalized agreement merely reflect interests of dominant powers. Legal rules then serve to bind weaker states in the system and enforcement relies on the willingness of dominant powers to bear the costs (Goldstein et al. 2000: 391). Furthermore, Downs et al. (1996) find that much of international agreements have very limited independent effect on state behavior, and mostly reflect what states would have done even in their absence (see Downs et al. 1996). Therefore, to the extent that states use power to secure preferable outcomes in negotiations, and to the extent that those outcomes reflect states' interests, the author of this thesis does not see a necessity to treat "legal" obligation as a distinct form of obligation that "[...] brings to play the established norms, procedures, and forms of discourse of international legal system" (Abbott et al. 2000: 411).

legalization (Abbott et al. 2000: 401–402). Of the three dimensions, obligation has the best utility for the purposes of the thesis.[39]

Several scholars have examined obligation features of arms control treaties (see Kreps 2018; Koremenos 2001; Abbott & Snidal 2000). Kreps (2018) suggests some ways to estimate degrees of obligation. First, high obligation is characterized by legally binding commitments and low obligation by political commitments. This difference can be illustrated on the Comprehensive Test Ban Treaty (CTBT), on one hand, and the Nuclear Suppliers Group, on the other. Second, high obligation does not allow for issuance of reservations, understandings and declarations, while low obligation does allow for such procedures. For example, the CTBT explicitly forbids the above mentioned, whereas the Treaty for the Prohibition of Nuclear Weapons in Latin America and the Caribbean allows issuance of understandings. Third, high obligation typically does not entail temporal conditionality, whereas low obligation does. For example, the Non-proliferation Treaty (NPT) entailed flexible trial period (Kreps 2018: 135–136).

Abbott and Snidal (2000) observe that most arms control agreements are of lower-obligation type, especially due to inclusion of the escape clause mechanisms. Indeed, various agreements include re-iteration of the same escape clause included in the Limited Test Ban Treaty (LTBT). Some other treaties such as the Outer Space Treaty also allow for withdrawal with an advanced notice. Generally, these various measures allow states to ameliorate distributional concerns, which are particularly severe when issues of security come to play (Abbott & Snidal 2000: 435). With the features of obligation identified, it is possible to conduct the conceptual-empirical synthesis (see Table 4). This allows connecting the conceptual basis with proposed regulatory solutions to AWS as described in the literature review. The five broadly defined categories—total prohibition, partial prohibition, hybrid regulation, minor regulation, and no regulation—reflect degrees of obligation with respect to criteria formulated by Kreps (2018).

[39] In addition, Abbott and Snidal (2000) argue that states often deliberately choose “soft law” over “hard law” due to its ability to lower certain costs of “hard law”. Their explanations include, for example, transaction costs, uncertainty, implications for national sovereignty, heterogeneity of preferences, and power asymmetries. They maintain that states opt for “hard law” as a way to increase credibility of commitments when noncompliance is difficult to detect, which is particularly true for arms control agreements (Abbott & Snidal 2000: 423).

Table 4: Empirical-Conceptual Synthesis of Forms of AWS Regulation

Label	Aspects of obligation	Examples of proposed solutions
Total prohibition	Strong legally-binding commitment; no reservations, understandings or declarations allowed; no temporal modifications allowed	"Full ban" or "pre-emptive ban", that is, blanket prohibition of the underlying technology through a treaty or additional protocol
Partial prohibition	Legally-binding commitment; reservations, understandings and declarations possible; temporal modifications possible	Treaty or protocol limiting use to specific targets such as non-humans, defensive purposes, interstate conflicts, and large battlefields, or prohibiting certain types of payloads, and specific stages of acquisition, R&D, testing, deployment, transfer, and proliferation, or establishing conditional use under MHC
Hybrid regulation	Weak legally-binding commitment; freedom of interpretation and temporal flexibility	Limited moratoria and framework conventions covering above mentioned issues subject to regulation
Minor regulation	Political rather than legally-binding commitment; significant freedom of interpretation and temporal flexibility	Resolutions, declarations, international and transnational dialogue, best practices, information-sharing agreements, codes of conduct, confidence-building measures to improve communication, transparency, and verification, covering above mentioned issues subject to regulation
No regulation	No obligations	No AWS-specific regulation, reliance on existing non-AWS-specific regulations, precedents, or technological solutions such as "ethical governor" software

Source: Author (2018)

4.3. Variables and Hypotheses: State Preferences and Institutional Activity

The thesis builds on several key assumptions to find which of the proposed regulatory solutions—ranging from no regulation to total prohibition—is likely to garner the most support from states involved in discussion on AWS. First, there is a danger of security dilemma and a potential for arms race due to various military advantages provided by AWS (see Chapter 2). States have incentives to build AWS, and the dilemma is exacerbated due to difficulty to differentiate between defensive and offensive systems, and the likely—albeit not demonstrated—advantage for the offense (see Altmann & Sauer 2017; Lachow 2017). To avoid the spiral of hostility, arms-racing, and thus reduce threats to national security, states signal their benign intentions by engaging in arms control discussions (see Glaser 1995). This is, to an extent, self-evident from the decision of majority of states to engage in discussions on AWS at international forums for conventional arms control, the First Committee at UNGA and CCW, in particular (see CSKR 2018a). Third, states use institutions, including those earlier mentioned, to further their interests, and they engage in the process of bargaining over distributional consequences of a potential arms control agreement that favors them the most, due to relative gains concerns (Fearon 1998: 273–274; Grieco 1988: 600).

Fourth, the potential agreement can take many forms with different degrees of obligation imposed on parties (see Goldstein et al. 2000). By the same token, thus proposed regulatory solutions mandate different degrees of obligation (see Chapter 3). Fifth, due to distributional concerns, countries tend to have different preferences as regards the specific form of such agreement, including the degree of obligation. Some seek to retain greater flexibility, others seek to bind the potential aggressors, and still others may refuse obligations due to greediness (see Abbott & Snidal 2000; Glaser 1995). Such heterogeneity of preferences results not only from external environment, most importantly the international anarchy, but also from domestic processes, responding to those external conditions (see Sterling-Folker 1997; Waltz 1979: 65, 122). One, and potentially the most significant factor, which explains the heterogeneity is the difference in technological level, because banning such an advanced technology as AWS benefits mainly the countries lagging behind in R&D (Hansel et al. 2018: 6). Sixth, agreements, and institutions generally, reflect interests of the powerful. Countries with the greatest share of power in the system therefore exercise the greatest influence over the resulting form of the potential arms control agreement (see Glaser 2003; Mearsheimer 1995; Jervis 1999). Power here is understood in accordance with neorealism primarily as material capability but it also accounts for "contextual" bargaining power (Waltz 1979: 131; Schelling 1956: 282). Following is the first set of hypotheses derived from this framework.

Hypothesis 1: States will prefer some form of regulation over no regulation at all.

Hypothesis 2: The most likely outcome of negotiations will reflect the interest of a state with the largest share of power.

With these fundamental assumptions described, it is important to elaborate on the first dependent variable—state preferences. More precisely, these are state preferences over the degree of obligation mandated by a potential AWS agreement. The thesis assumes that state preferences can be inferred from their statements at the aforementioned international forums. However, any such endeavor faces some hurdles. As Jervis (1988) suggests: "Actors rarely give complete statements of their preferences [...]" (Jervis 1988: 323). As the CSKR movement demonstrates, certain states might be very vocal and relatively clear about their preferences, while most states reserve judgement on a specific regulatory solution, retaining a wait-and-see posture (Jenks & Ford 2016; Docherty 2014). The only way to systematically infer state preferences is, thus, to estimate them indirectly according to a priori criteria, which have been elaborated in the preceding section (see Table 3). Next, not only is complete information on preferences unattainable but, especially in a bargaining situation, states often give incomplete or misleading statements on purpose so as to achieve their most preferred outcome (Schelling 1958: 282). Consequently, the dependent variable captures instead what states "want" others to know. Of course, under neorealist assumptions, states ascribe utmost preference to their survival in the system (Waltz 1979: 126). When negotiating on the backdrop of the potential arms race, preferences should thus, at minimum, reflect states' security interest, which, in turn, inform the decision on arms control.

In addition to preferences, the thesis employs the concept of state power as an independent variable to test the first two hypotheses. As previously mentioned, the thesis employs the concept of "hard" power, and, particularly, power based on material resources. Nye (2011) notes: "A pragmatic or common-sense realist takes into account the full spectrum of power resources, including ideas, persuasion, and attraction" (Nye 2011: 18). Facets of power such as framing and agenda-setting are certainly an important feature of the current discussions, beside the typical realist power of coercion and payment (see Nye 2011). The author of this thesis does not seek to deny these alternative resources of state power. Nonetheless, attempting at measuring intangible factors in a debate over material capability may be ultimately ill-advised. An arms control agreement without participation of states with both an incentive to develop AWS and capacities—that is, hard power—to do so, would be a hollow agreement. Additionally, the greater are the incentives and capacities the lower is the cost of no agreement, and thus greater leeway in negotiations (Fearon 1998: 279). Thus, arms control can be successful only when the agreement reflects preferences of the most powerful states. For sake of comparison, the thesis adopts additional measure of "smart" power to examine how the results would differ (see Nye 2011).

Furthermore, the thesis employs the aforementioned assumptions in order to find, what explains the activity of states at various forums, including HRC, UNGA, and CCW, where the discussions on AWS take place. Specifically, the third and fifth assumptions suggest that states' institutional activity should reflect the stakes they have in preventive arms control. In turn, states' institutional activity should reflect their interest in the topic, but more specifically, also the capacities and incentives engage in AWS-related R&D. Thus, the primary motives of states involved in AWS debates, including those rallying for the cause of CSKR, are by no means altruistic. The author assumes that states most invested in R&D of AWS-related technologies and countries with positive incentives and capacities to procure advanced military systems have a greater stake in the discussion, especially if it leads to unfavorable distributional arrangements in favor of those that lack the aforementioned (Hansel et al. 2018: 6). Accordingly, the author further assumes that states use institutions, in this case HRC, UNGA and CCW forums, to assert their interest, ensuring that any potential distributional consequences do not affect them disproportionately. Most importantly, defensive realism suggests that states can use institutions to alleviate the security dilemma and arms-racing tendencies through signaling of their benign motives (Glaser 2003: 409). This latter function of international institutions seems particularly useful to countries that both fear other' intentions with respect to R&D of AWS-related technologies and "are feared" by others.[40]

To explore, which capacities and incentives in particular lead countries to participate more actively in AWS discussions, the thesis turns to literature on arms racing and arms trade, specifically. First, the literature mentions external threats, or *Conflict*, as one of the most influential drivers of arms buildup (Glaser 2000: 251; Buzan & Herring 1998: 83; Pearson 1989: 154). Advanced economies with high income, *GDP per capita*, have greater capacities to buy and develop advanced technologies including those related to UVs and AWS R&D (Fuhrmann & Horowitz 2017a: 410; Martínez-Zarzoso & Johannsen 2017: 7; Pearson 1989: 155). Similarly, *R&D spending* indicates nation's capacity to develop high-tech systems, as well as the interests of a domestic R&D complex involved in development of autonomous systems (Glaser 2000: 251, 257; Buzan & Herring 1998: 83). *Military expenditure* is the key driver of weapons procurement specifically (Pamp & Thurner 2017: 460–461; Pearson 1989: 154–155). The character of the regime also affects its motives and incentives. AWS are especially attractive to *Democracy* due to public societies' sensitivity to casualties, or the body-bags returning back home, which often affects the support to nation's military operations. This is because AWS like drones allow removing soldiers the battlefield (Arkin 2017: 37; Fuhrmann & Horowitz 2017a: 415; US Army 2017: 2; Marchant et al. 275). Countries' military ambition and their desire to project power and influence on an international level are also reflected in

[40] Of course, this assumption applies to security-seekers but not so much to "greedy" states that might not see such a value in engaging in discussions, let alone discussions about arms control, especially due to the likely advantage of offensive AWS capabilities and due to indistinguishability of offense and defense. Furthermore, even if a greedy state participates actively in discussions on AWS, this does not preclude the possibility of bluffing and later cheating on the agreement (Glaser 1995: 65–69). These risks are also exacerbated by the many difficulties in effective verification of AWS, in particular, due to their dual-use nature and the ability to, essentially, change an offensive system to a defensive one through software updates (Wallach 2017b: 29).

the degree of *Militarization* of society (Martínez-Zarzoso 2017: 7). Lastly, existence of a *UCAV program* in a given country should in itself provide a powerful incentive because UCAVs are a crucial precursor to future AWS (Altmann & Sauer 2017: 118, 123). Following is a set of hypotheses derived from the incentives and capacities.

Hypothesis 3: *Institutional activity* will correlate positively with *Conflict*.

Hypothesis 4: *Institutional activity* will correlate positively with *GDP per capita*.

Hypothesis 5: *Institutional activity* will correlate positively with *R&D spending*.

Hypothesis 6: *Institutional activity* will correlate positively with *Military expenditure*.

Hypothesis 7: *Institutional activity* will correlate positively with *Democracy*.

Hypothesis 8: *Institutional activity* will correlate positively with *Militarization*.

Hypothesis 9: *Institutional activity* will correlate positively with *UCAV program*.

5. Methodology

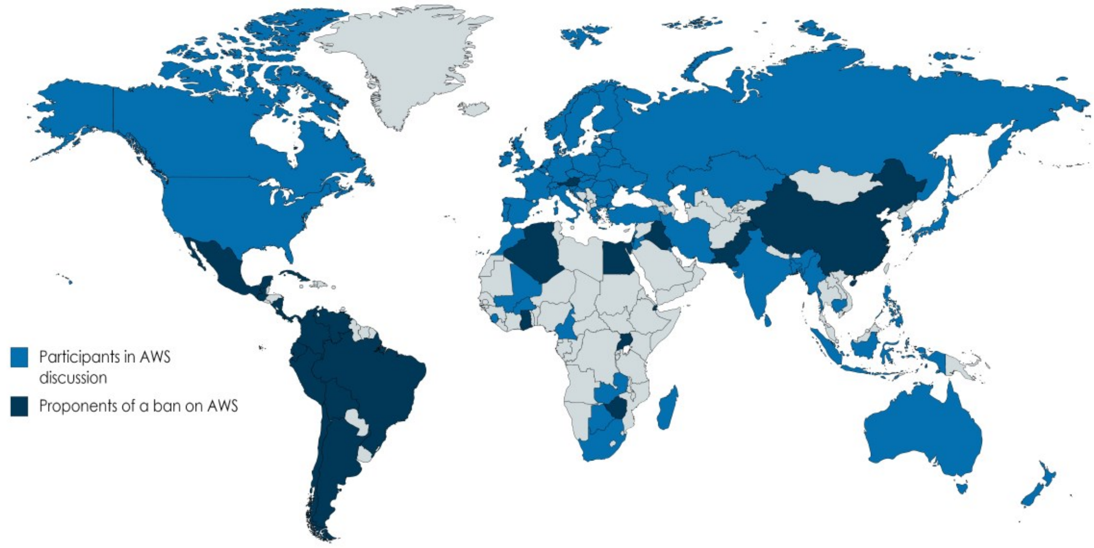
The purpose of this chapter is to provide a research design for testing of the hypotheses and answering the two research questions. In so-doing the thesis thus elaborates on the rationale behind case selection, operationalization of variables, methods used for analysis, and data. After exploring the reasons for case selection, which covers 12 most technologically advanced and powerful states in discussion on AWS—namely, the US, UK, China, Russia, India, Brazil, Japan, Korea, Israel, Italy, France and Germany—the thesis continues by addressing the respective variables, methods and data, separately for each research question. This is important so as to maintain clarity with respect to the research design. Accordingly, the two research questions and associated hypotheses will be addressed in separate subchapters in the following chapter (see Chapter 6). This chapter will, first, address the variable *State preferences*, describing its operationalization through content analysis of the statements delivered at the various multilateral forums and the subsequent use of spatial analysis by constructing a policy continuum. The MVT model for prediction of the most likely outcome is described as well as the different measures of power used for prediction (see Bueno de Mesquita 2013). Second, the chapter elaborates on the method regression analysis and relevant diagnostics, operationalization of *Institutional activity* and other variables (see Monogan 2015; Allison 1999).

5.1. Case Selection: The Technologically Advanced and Powerful States

To answer the first research question, the thesis focuses on statements from a limited number of countries during the period from 2013 to 2017. First, it should be clarified what kind of statements? As was earlier mentioned, the current debate takes place mainly at three multilateral forums, including HRC, First Committee of UNGA, and CCW. The first debate on AWS took place in 2013 at HRC in response to the report presented by Christof Heyns (2013). Parallel debates then started at the UNGA and CCW. The CCW, in particular, has become the center of state discussions, especially since the commencement of formal meetings of experts on AWS, as well as due to the recent establishment of the Group of Governmental Experts (GGE) on AWS. Apart from state delegations, multiple NGOs—for example, the International Committee for Robot Arms Control, HRW, Article 36, PAX, Mines Action Canada, or others—under the banner of the CSKR also regularly participate in the meetings at these forums. Up until this point, 85 countries have expressed their views on AWS, and 26 of them have expressed their support of a ban on fully autonomous weapon systems, with China recently joining this group (CSKR 2018a; CSKR 2018b). A map on the following page illustrates the geographical representation of participants in debates on AWS (see Figure 10). The darker color stands for proponents of a ban, whereas the lighter one for the rest of the participants. Importantly, states express their views at the aforementioned fora through their national delegations, which, in turn, deliver public statements, position papers, and other documents available online after the meeting (CSKR 2018a; Reaching Critical Will 2018; UNOG 2018). It is precisely these documents, or "statements", from which the thesis seeks deriving state preferences.

Second, it should be clarified what countries? As was mentioned, the thesis does not seek to cover the views of all participants in discussions. This is, to a large extent, due to time constraints associated with data collection, and because most of the participants have delivered only a handful of statements. Obviously, attempting to estimate states' preferences from a limited number of observations would be ill-advised with respect to accuracy of the estimation. To find, which of the outcomes of the discussion is most likely, the thesis thus focuses on only several of the most powerful states, both as regards the distribution of material capabilities as well as the context of bargaining over the distribution of costs and benefits. Importantly, if a legally binding and effective agreement were to be concluded, this will ultimately require ratification by states leading in AWS R&D, and other related fields such as military robotics (Dunkelberg 2016: 8–9). These include, in particular, the US, UK, Russia, France, Germany, Italy, Japan, India, Israel, S. Korea, and China (Boulanin & Verbruggen 2017: 94; Roff 2016; Scharre & Horowitz 2015: 21–23). The thesis is going to examine the same set of countries with addition of Brazil. This will enable capturing top 9 countries in terms of the share of hard as well as smart power (see Nye 2011). Furthermore, including Brazil will also hopefully diversify preferences of AWS regulation as Brazil voiced its support to a ban (CSKR 2018b). The author also considered several other criteria for case selection (see Table 5).

Figure 10: Participants of the Debate on AWS[41]



Source: Author

[41] These are, namely: Algeria, Argentina, Australia, Austria, Bangladesh, Belarus, Belgium, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Cambodia, Cameroon, Canada, Chile, China, Colombia, Costa Rica, Croatia, Cuba, Czech Republic, Denmark, Djibouti, Ecuador, Egypt, Estonia, Finland, France, Germany, Ghana, Greece, Guatemala, Holy See, Hungary, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Japan, Jordan, Kazakhstan, Kuwait, Latvia, Lebanon, Lithuania, Madagascar, Mali, Mexico, Moldova, Montenegro, Morocco, Myanmar, Netherlands, New Zealand, Nicaragua, Norway, Pakistan, Palestine, Panama, Peru, Philippines, Poland, Portugal, Romania, Russia, Sierra Leone, Slovakia, Slovenia, South Africa, South Korea, Spain, Sri Lanka, Sweden, Switzerland, Turkey, Ukraine, Uganda, United Kingdom, United States, Venezuela, Zambia, Zimbabwe (CSKR 2018b).

Table 5: Summary Statistics on Countries of Case Selection

Country	Hard power ranking as of 2013[42]	Smart power ranking as of 2017[43]	Share of arms exports 2013–2017[44]	R&D spending in $ b. as of 2015[45]	High-technology exports in $ b. as of 2016[46]	UCAV program as of 2017[47]	Support for a ban as of 2017
China	1st	2nd	5.7 %	15463.3	560.1	Yes	No
US	2nd	1st	34 %	43302.8	148.5	Yes	No
India	3rd	4th	0.1 %	1444.5	16.7	Yes	No
Russia	4th	3rd	22 %	1788.9	8.7	Yes	No
Japan	5th	8th	0.1 %	19546.0	105.1	No	No
S. Korea	6th	16th	1.2 %	4953.2	130.5	Yes	No
Brazil	7th	9th	0.2 %	2890.0	8.4	No	Yes
Germany	8th	5th	5.8 %	10077.5	193.8	Yes	No
UK	9th	6th	4.8 %	4256.4	69.2	Yes	No
France	10th	7th	6.7 %	6072.3	113.3	Yes	No
Italy	15th	12th	2.5 %	2676.9	29.7	Yes	No
Israel	45th	28th	2.9 %	1074.7	9.6	Yes	No

Source: Author (2018)

42 Data come from Mijares (2016), and cover only the period until 2013 (see Mijares 2016). The measure is described in detail in the following section.
43 Data come from In.europa (2017), and cover only the year 2017 (see In.europa 2017). The measure is described in detail in the following section.
44 Data come from Wezeman et al. (2018), and cover the period 2013–2017 (see Wezeman et al. 2018).
45 Data come from the World Bank database (see WB 2018d).
46 Data come from the World Bank database (see WB 2018f).
47 Data come from Fuhrmann and Horowitz (2017) (see Fuhrmann & Horowitz 2017b).

5.2. Methods of Analysis: Content Analysis and Median Voter Prediction

The thesis adopts a variety of quantitative research methods to answer the two research questions and relevant hypotheses. As for the first research question, the thesis adopts a one-dimensional spatial model based on median voter theorem (MVT). Spatial models of political competition represent a well-established means of analysis based on the work of Duncan Black (1958) and Anthony Downs (1957) (see Black 1958; Downs 1957). The assumption is that political actors and policies can be placed in a political "space" using Euclidean geometry (Dowding 2011: 688; Laver 2003: 6). In addition, it is assumed that actors have an "ideal point" or most-preferred policy that can be compared to other points in the space in terms of the relative distance (Laver 2003: 4). MVT constitutes one such method, which allows comparison of actors' positions—or in this case, states' preferences towards AWS regulation. The theorem says: "[...] if issues are one-dimensional, preferences are single-peaked [...] it takes a majority to win, and everyone with a vote votes based on their preferences then the median voter's position is the winning outcome" (Bueno de Mesquita 2013: 102). In other words, if this holds, the position in the "middle" is the winning outcome.[48]

Bueno de Mesquita (2013) suggests particular adjustments in the context of international politics. This involves substituting the rule of majority of votes with majority of power. Thereby reformulated MVT tells us that the first position on the dimension which gains the majority of power is the winner. More precisely, the actor whose position is associated with the majority of power secures her most-preferred outcome (Bueno de Mesquita 2013: 110).[49] Of course, this conception of international policy-making is very much in line with neorealist arguments about

[48] The assumption of unidimensionality states that the issue in question must be meaningfully displayable on a straight line. The assumption of single-peaked preferences states that all choices on both ends of the dimension must be less preferred compared to the ideal point. In addition, based on the distance from ideal point, it is possible to compare alternatives and determine which of the points is preferable. Finally, the assumption of majority rule states that if majority is required to win, then the position in the middle—meaning that half of the voters are located to its right and left—is the winning outcome (Bueno de Mesquita 2013: 102–114; Morrow 1994b: 108–109)

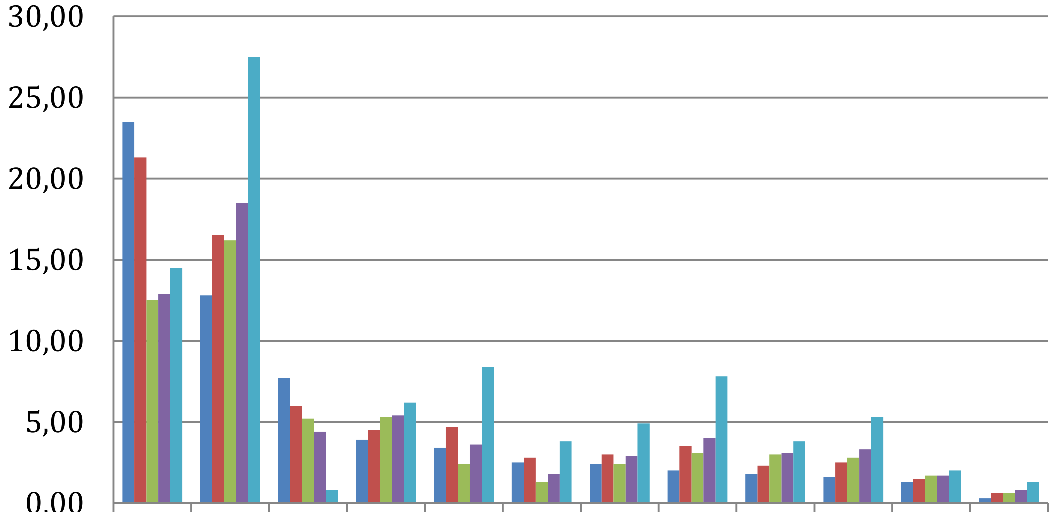
[49] To find the median voter can be found using a graph of cumulative total amount of power, or the votes, plotted against the policy continuum, which can be divided into tenths, or in our case, fifths, of a point swaths associated with particular policy positions. The point where the amount of power hits the 50% mark is the median outcome (Bueno de Mesquita 2013: 111–112).

international institutions reflecting the interests of the powerful (see Glaser 2003; Mearsheimer 1995; Jervis 1999). According to Bueno de Mesquita, power depends on the context, or the policy dimension, in question, and refers to relational power to coerce, persuade, and else to make the others adopt one's most preferred option (Bueno de Mesquita 2013: 110). The thesis operationalizes two concepts of power, including through the Composite Index of National Capability (CINC), reflecting the material foundations of hard power, and State Power Index (SPI), corresponding to the concept of "smart" power (see In.europa 2017; Mijares 2016). Additionally, the role of context is accounted for in modified versions of the two measures based on countries' industrial base for AWS development.

First, CINC constitutes a measure of state power designed by J. David Singer (1963), using the percentages of world totals within six categories. These include, in particular, demographic, economic, and military indicators (Mijares 2016: 1–2). The latest update of the measure comes from Mijares (2016), and covers the years 2008 to 2013 (see Mijares 2016). Second, SPI constitutes a slightly more complex measure, which accounts for factors such as economic capital, militarization, land, human resources, culture, natural resources, and diplomacy, with different weight assigned to each factor. Data on SPI cover only the year 2017 (In.europa 2017). As regards the role of the context—for example, issue salience and stakes for a given state—the thesis employs an AWS Industrial Base (AWSIB) index, which captures percentages of world totals in R&D spending, arms exports, and high-technology exports. AWSIB was designed by the author to encompass countries' capacities to develop AWS. At the same time, the measure accounts for findings of contributions on drivers of development of military robotics (see Boulanin & Verbruggen 2017; Fuhrmann & Horowitz 2017a; Scharre & Horowitz 2015). Adjustment of the initial measurements of CINC and SPI is subsequently achieved simply by adding AWSIB index into respective equations. Results of this modification, including the original measurements can be seen in the following graph (see Figure 11). The author uses labels CCINC and CSPI to distinguish modified measurements.[50]

[50] See Appendix 2 for formulas on CINC, SPI, AWSIB, and modified version of CINC (CCINC) and modified version of SPI (CSPI).

Figure 11: Distribution of Power for Countries of Case Selection (%)[51]



	CHN	USA	IND	RUS	JPN	KOR	BRA	DEU	GBR	FRA	ITA	ISR
CINC	23,50	12,80	7,70	3,90	3,40	2,50	2,40	2,00	1,80	1,60	1,30	0,30
CCINC	21,30	16,50	6,00	4,50	4,70	2,80	3,00	3,50	2,30	2,50	1,50	0,60
SPI	12,50	16,20	5,20	5,30	2,40	1,30	2,40	3,10	3,00	2,80	1,70	0,60
CSPI	12,90	18,50	4,40	5,40	3,60	1,80	2,90	4,00	3,10	3,30	1,70	0,80
AWSIB	14,50	27,50	0,80	6,20	8,40	3,80	4,90	7,80	3,80	5,30	2,00	1,30

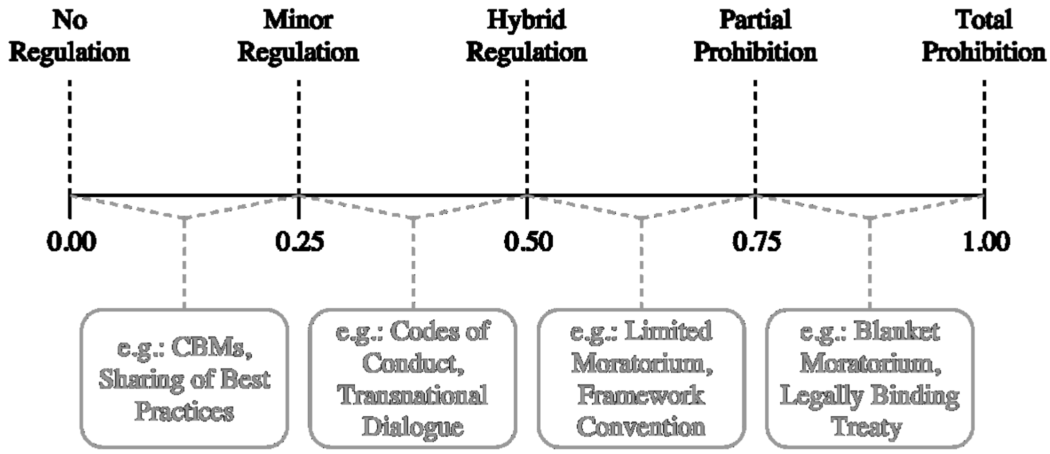
Source: Author (2018)

In addition to state power, it is important to identify the policy dimension in question, and ensure that the array of policy options can be meaningfully displayed on a straight line (Bueno de Mesquita 2013: 102). As described earlier, when states engage in arms control negotiations, regulatory options they can choose to enforce vary in terms of the degree of legalization, including the aspect of obligation (see Kreps 2018; Abbott et al. 2000; Goldstein et al. 2000). Moreover, as suggested in the empirical-conceptual synthesis of forms of AWS regulation (see Table 4), the degrees of obligation can be associated to specific solutions to AWS, which were discussed in the literature review (see Chapter 3). Thus defined policy continuum can be illustrated on the following figure (see Figure 12). With a policy continuum in place, this enables placing actors involved in the political competition in the Euclidean space, comparing preferences to other available options by calculating

[51] It should be noted that the graph here shows countries' share of world totals, rather than share for a given case selection. Furthermore, totals account for 62.7%, 69.2%, 56.5%, 62.4% and 78.6% respectively for CINC, CCINC, SPI, CSPI and AWSIB indexes.

distances between their ideal points and alternative policies, and distances to others' idea points (Bueno de Mesquita 2013: 105–109). However, this requires estimating positions of actors in advance (Slapin & Proksch 2008: 705–706).[52]

Figure 12: AWS Policy Continuum[53]



Source: Author (2018)

To procure data and estimate state preferences from statements by national delegations, the thesis uses the method of content analysis (see Krippendorff 2013; Neuendorf 2000). Within the tradition of quantitative research, content analysis is used primarily as a method for data collection (Schreier 2014: 173). Krippendorff (2013) defines the method as "[...] a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (Krippendorff 2013: 24). In practice, this technique involves working through the texts and registering occurrence of predefined categories based on a codebook and a coding form (Marying 2014: 12). Neuendorf (2002) identifies key principles

[52] In comparative politics, methods for estimating actors' positions include, for example, expert surveys, public opinion polls, and hand-coded or computer-based content analysis of manifestos, speeches, and statements, as a means of data collection (Slapin & Proksch 2008: 705–706; Benoit & Laver 2006: 57–58; Mair 2003: 13–15).

[53] Importantly, policy options displayed on the continuum present ideal-cases. Labels below the continuum provide examples of specific solutions but they should be seen as illustrative, rather than given.

of content analysis as: an a priori design, reliability, validity, generalizability, and replicability (Neuendorf 2002: 10–13). In line with these principles, the author has created an a priori design, including definition of each category, measurement, and coding rules, and a corresponding codebook (see Table 6) and coding form. States' preferences are assessed within four broad categories of issue-areas, which reflect the empirical-conceptual synthesis of issues in AWS regulation (see Table 3). The data, or the statements, come from the websites of a project Reaching Critical Will, which provides an archive of states' statements from meetings at the UN, UN Office at Geneva, which provides data on meetings at CCW, and the CSKR, which provides data on the HRC meetings (CSKR 2108a; UNOG 2018; Reaching Critical Will 2018).

Each statement—that is, the input document submitted by the delegation—is divided in individual paragraphs, which constitute units of observation.[54] This is done so as to make the estimation more accurate, and, by extension, reliable and valid rather than categorizing the whole document, which often consists of several pages. Furthermore, the length of the documents can vary from several sentences to multiple pages. Observations on the level of paragraphs, on the other hand, will enable capturing a coherent thought on a specific issue. This adjustment brings us to almost 700 units of observation that address AWS directly or indirectly. The variable *State preference* is subsequently calculated on the level of unit of analysis, meaning on the level of individual countries, as a sum of scores for each category divided by four. Scores for individual categories at the level of unit of analysis are also calculated as a sum of values of all observations divided by the number of the observations. Thus obtained scores of *State preference* ranges between 0 and 1 and allows placing states on the policy continuum, and testing the first hypothesis by applying the MVT. Results of the content analysis will be presented in the following chapter (see Chapter 6), using frequency tables and visually in word clouds.[55]

[54] For the difference between unit of analysis and unit of observations, see Neuendorf 2002: 13.
[55] Word clouds are "[...] a method of visually representing text data [...] typically used as a summative way to display the frequency of keywords [...]" or "[...] providing a visual of the most important topics [...]" (Ramlo 2011: 100–101). By the same token, the visuals here can help discern some of the key topics brought up in states' statements, providing a picture of the overall trajectory of the debate on AWS, and enhancing clarity (see Dickinson 2010).

Table 6: Content Analysis Codebook

<table>
<tr><th rowspan="2">Category</th><th rowspan="2">Coded issues</th><th colspan="4">Assessment of State Preferences</th></tr>
<tr><th>Positive</th><th>Neutral</th><th>Negative</th><th>Absent</th></tr>
<tr><td>International law</td><td>Issues related to adherence to international law, including other arms control agreements, potential and existing forms of regulation</td><td rowspan="4">Acknowledgement of specific issues; explicit judgement in favor of AWS’ ability to cope with specific issues or explicit judgement against necessity for corrective and preventive measures to address specific issues or explicit judgement in favor of AWS’ ability to improve upon specific issues; clear or coherent view on specific issues</td><td rowspan="4">Mere acknowledgement of potential issues; absence of explicit judgement on specific issues; conflicted or incoherent view on specific issues</td><td rowspan="4">Acknowledgement of specific issues; explicit judgement against AWS’ ability to cope with specific issues or explicit judgement favoring necessity for corrective and preventive measures to address specific issues or explicit judgement against AWS’ ability to improve upon specific issues; clear or coherent view on specific issues</td><td rowspan="4">Lacking acknowledgement of issues; absence of explicit judgement on specific issues; incomplete view on specific issues</td></tr>
<tr><td>International security</td><td>Issues related to international security such as arms racing, strategic stability and proliferation, as well as military power, strategy and tactics</td></tr>
<tr><td>Societal security</td><td>Issues related to protection of humans and society, including political and societal systems as well as value systems, ethics and morality</td></tr>
<tr><td>Operational safety</td><td>Issues related to technical performance and operational safety, including those stemming from complexity, predictability, reliability, and vulnerability</td></tr>
<tr><td colspan="2">Coded as</td><td>0</td><td>0.5</td><td>1</td><td>NA</td></tr>
</table>

Source: Author (2018)

Before moving on, it is important to describe multiple procedures and rules the author has adopted to make the estimation more accurate. First, a paragraph is hereby defined as a text consisting of at least 2 sentences. Thus, when a paragraph in the document submitted by a country contains only one sentence, it is added to the previous or the following paragraph based on the specific context. As a general rule, when the previous and the following paragraph are thematically identical, the isolated sentences are added to the shorter paragraph. Second, coded issue-areas can overlap. Therefore, a paragraph can refer to issues related to international law, international security, societal security, and operational safety at the same time, if a state in question expresses its opinion on all of those simultaneously. Third, if a statement refers directly to particular solutions to AWS regulation, coding reflects the suggested categorization provided in Table 4. To give an example, when a state calls for a ban or prohibition on AWS, its corresponding statement is always coded as negative (1) in the issue-area of international law. Conversely, when a state calls for sharing of best practices, its statement is always coded as positive (0).[56]

5.3. Methods of Analysis: Regression Analysis

To answer the second research question and relevant hypotheses, the thesis employs a linear regression model. This is perhaps the most widely used statistical method of analysis in Political Science (Monogan 2015: 79; Allison 1999: 1–2). The method is used to explore the relationship between the dependent variable, in this case *Institutional activity*, and one or more independent variables. More precisely, it allows studying what causes change in the dependent variable, in what direction does the change occur and what is the magnitude of a given independent variable's effect on the dependent variable. There is a plethora of variants of the method but the most commonly used and the one used in this thesis is "ordinary least squares multiple linear regression," or OLS (Allison 1999: 2). The model can be expressed as $y_i = B_0 + B_1 x_{i1} + B_2 x_{i2} + \ldots + B_p x_{ip} + E$, where i stands for observations, y_i for

[56] The latter adjustment is based on the assumption that when a state gives its opinion on a specific regulatory solution, we can directly infer its position on the degree of legalization, rather than inferring it indirectly.

the dependent variable, $x_{i1}...x_{ip}$ for independent variables, E for residuals or the random error in regression prediction, B_0 for the intercept, and, lastly, $B_1...B_p$ for regression coefficients (Investopedia 2018). Whether an independent variable has an effect on the dependent one can be found by calculating *p value* associated with regression coefficients. In general, where the *p value* is less than 0.05 we conclude that the effect of the variable is significant or, more precisely, that it does not equal zero (Allison 1999: 15–16). Furthermore, coefficients can tell us in what direction, whether negative or positive, does the change occur, and what is the magnitude of this change (Allison 1999: 27). Additionally, we can calculate the measure of R^2 to find how much variance in the dependent variable does the model explain (Allison 1999: 31).

In order to ensure that the results of the regression analysis are not biased, we have to test whether the model satisfies the regression assumptions. These are, specifically, the assumption that predictors x are independent of the random error, conditional mean of the error is equal to zero, residuals have constant variance, or, in other words, we observe "homoscedasticity" as opposed to "heteroscedasticity", there is no autocorrelation between random errors, and they should have a normal distribution (Monogan 2015: 85; Allison 1999: 122–123). If these conditions hold, the model is unbiased, efficient, and our inferences are valid (Monogan 2015: 85). First, an essential method of regression diagnostics is to plot the residuals against fitted values and the independent variables. When the scatterplot shows pattern of rising and falling, we need to reconsider the functional form, and consider models other than the standard OLS model, for example a generalized least squares model. To test whether heteroscedasticity is, indeed, present, we can calculate a Breusch-Pagan score. When the *p value* of the test is below 0.05, there is heteroscedasticity and results should be corrected. One way to correct the results is to use the robust standard errors. In addition, to test whether residuals are normally distributed, we can use quantile-quantile plots or calculate the Jarque-Bera statistic, with *p values* below 0.05 indicating that residuals are not normally distributed. Finally, we have to test for multicollinearity—that is, whether the effect of an independent variable does not result from correlation with other ones. To do this, we can calculate "VIF" scores with values below 10 indicating no sever issues (Monogan 2015: 85–93).

The dependent variable *Institutional activity* was constructed by the author based on states' attendance to the meetings of the HRC, UNGA First Committee and CCW. More specifically, the value of *Institutional activity* represents the number of meetings where a state has delivered a statement or statements mentioning AWS. The resulting number is thus an interval variable with a minimum of zero in case of countries which have not yet publicly addressed the issue. Thus operationalized *Institutional activity* also arguably works as an approximation of states' interest in the topic. In other words, the more incentives and capacities a country has in AWS-related R&D, the more interested it is in the topic, and the more actively it should participate in the meetings. The data on *Institutional activity* come from websites of a project Reaching Critical Will, which provides an archive of states' statements from meetings at the UN, UN Office at Geneva, which provides data on meetings at CCW, and the CSKR, which provides data on the meetings at the HRC (CSKR 2108a; UNOG 2018; Reaching Critical Will 2018). The author has calculated values for the dependent variable as a sum of country's attendances throughout the years 2013-2017. The data cover 155 countries.[57]

As for the independent variables, multiple measures come from the World Bank database. First, *GDP per capita* is an interval variable covering the years 2013 to 2017 and 144 countries in the dataset. It provides information on states' level of economic development in annual USD (WB 2018a). Next, variable *Militarization* is calculated as military personnel per capita, an interval variable, from data on total armed forces personnel and total population. The data cover the period from 2013 to 2016 and 148 countries in the dataset (WB 2018b; 2018c). Finally, data on *R&D spending*, again, an interval variable, were procured using the data on research and development expenditure as a share of GDP and total GDP as a total amount spent on R&D in annual USD and cover the years 2013 to 2015 and 96 countries in the dataset (WB 2018d; 2018e). Next, two variables come from the Center for Systemic Peace database. The variable *Conflict* comes from the dataset on Major Episodes of Political Violence and constitutes a categorical variable indicating the magnitude of both inter- and intra-state conflicts from 2013 to 2017 for 152 states in our dataset

[57] It should be noted that several countries were left out of analysis because the author was unable to find, whether the country mentions the issue of AWS in its statement. This is due to the fact that several countries submit their statements only in Arabic, and no translation was publicly available.

(Center for Systemic Peace 2018a). The variable *Democracy* comes from the Polity IV dataset, constitutes a categorical variable. It indicates the total polity score that ranges from most autocratic to most democratic regimes, and covers the period of 2013 to 2017 and 148 countries in our dataset (Center for Systemic Peace 2018b). Furthermore, the interval variable *Military expenditure* comes from SIPRI database and indicates expenditure in annual USD, covering the years 2013 to 2017 and 147 countries in the dataset (SIPRI 2018). The *UCAV program* is a dummy variable that comes from replication dataset of a study by Fuhrmann and Horowitz (2017a), and indicates existence of a program of armed drones as of the year 2014. To author's best knowledge, as for now, there are no more recent data (Fuhrmann & Horowitz 2017b).

In addition, the thesis includes two control variables, both dummy, *LDC* and *Ban*. First, the purpose of including the variable LDC is to ensure that the variation in *Institutional activity* does not result merely from the difference in general levels of development, and that we are actually measuring the effect of specific capacities and incentives associated with AWS-related R&D. Information on "least developed country" status come from the UNCTAD (UNCTAD 2018). Second, the variable *Ban* is supposed to control for *Institutional activity* of countries that have affiliation to the CSKR movement. The assumption here is that supporters of a ban on AWS, like those rallying against it, should have pronounced distributional concerns, and thus greater interest in the discussions on average. The dummy further ensures that the change in *Institutional activity* does not depend merely on the political-ideological activism of certain states. Information on ban supporters as of the end of the year 2017 comes from the CSKR website (CSKR 2018b). Descriptive statistics for all of the above mentioned variables are to be found in Table 7 on the following page. In addition, it should be noted that the author uses "r-studio" software to conduct the statistical analysis.

Table 7: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Min
Institutional Activity	155	2.471	3.850	0	14
Conflict	152	0.428	1.285	0	7
GDP per capita	144	13892.52	20298.88	255	116612.9
UCAV program	155	0.168	0.375	0	1
Democracy	148	15.642	5.904	1	21
Ban	155	0.142	0.350	0	1
R&D spending	96	1426494	5359723	49.776	44846035
Military expenditure	147	8641209	46495055	0	552568000
Militarization	148	0.005	0.007	0	0.055
LDCs	155	0.265	0.443	0	1

Source: Author (2018)

6. Analysis

As was already mentioned, the analysis proceeds in two subchapters, using the variables and methods for testing the relevant hypotheses. The purpose of the chapter is, first, to analyze the substantive content of states' statements at the HRC, UNGA and CCW, providing an extensive empirical account of AWS discussions with regard to legal, ethical, security and operational aspects, estimating countries' ideal points and placing them on the policy continuum. Additionally, the thesis explores the content through a word cloud and n-grams of common phrases which serve, in part, to legitimize the use of statements as the primary data source. Subsequently, model of median voter prediction is applied with several adjustments as described by Bueno de Mesquita (2013) to estimate the most likely outcome based on power associated with different policy positions. Second, the chapter employs a model of linear regression to explore the relationship between *Institutional activity* and *GDP per capita*, *R&D spending*, *Conflict*, *Democracy*, *Military expenditure*—the capacities and incentives associated with AWS-related R&D. Finally, methods of regression diagnostics are employed to examine, whether the presented models are effective and unbiased. Based on the results of the analysis, the thesis provides an answer to each question, while confirming the vast majority of hypotheses.

6.1. Substance of Discussions on AWS: State Preferences and the Likely Outcome

Upon conducting a content analysis of some 700 paragraphs, the author has identified a total of 433 neutral, 59 negative and 172 positive statements related to the four previously described issue-areas. In the light of these findings, the author has decided to make one final adjustment to improve the accuracy of estimation of state preference. This is because the high number of cases coded as neutral (67%) could distort the estimation. Unlike positive and negative statements, neutral ones merely acknowledge the possibility of certain issues. In practice, this means that a state often invokes certain issues without necessarily judging their severity. As an example, delegations often say that AWS are "[...] a highly complicated matter in its legal, political, and ethical aspects" (Chinese Delegation 2013). While the statement satisfies the criteria for coding as a neutral statement in the area of international law and societal security, it says little to nothing about a given state's preference of AWS regulation. For this reason, the thesis applies double weight to both negative and positive observations, in part also to ensure greater variance in preferences of states involved in the discussions.

Figure 13: Word Cloud of the Debate on AWS[58]

Source: Author (2018)

With respect to the aim of providing an extensive empirical account of the overall trajectory of the debate, the input documents of the content analysis allow identifying some of the key themes. The word cloud above (see Figure 13) contains all statements. At the most general level of analysis, the word cloud illustrates that

[58] Data for the word cloud come from all statements analyzed in the content analysis. Furthermore, the threshold for visualization was set to maximum of 300 words. Size indicates word frequency but the purpose of using different colors is purely for greater clarity, so that words can be clearly distinguished from one another.

the most common words in states' statements include, for example: "autonomous", "weapons", "systems", "lethal", "LAWS", "international", "discussions" and various others. With respect to specific issues within the four previously described issue-areas, words such as "operational", "technical", "military", "ethical", "humanitarian" or "legal" seem to be important, although they appear much less prominently than the aforementioned. This observation generally legitimizes the thesis' focus on the discussions at UNGA, HRC, and CCW as a substantive reflection of states' opinions on AWS, including in the four broad issue-areas. Perhaps a more accurate analysis of the key themes can be provided by looking at the most common n-grams—that is, words that follow others immediately (Silge & Robinson 2018). The table below suggests that the topical focus of discussions lies mainly in the legal domain.

Table 8: Selected N-Grams of Discussions on AWS[59]

Bigrams			Trigrams		
Phrase	Count	Freq.	Phrase	Count	Freq.
Weapon systems	227	0.41	Autonomous weapons systems	160	0.28
Autonomous weapons	221	0.39	Lethal Autonomous Weapons	146	0.26
Lethal autonomous	170	0.29	International humanitarian law	90	0.16
International law	133	0.24	Meeting of Experts	64	0.11
Humanitarian law	91	0.17	A common understanding	32	0.06
International humanitarian	91	0.17	Development and use	32	0.06
Human control	61	0.11	High Contracting Parties	31	0.06
Article 36	54	0.10	Definition of LAWS	28	0.05
Common understanding	50	0.09	Certain Conventional Weapons	23	0.04
Emerging technologies	43	0.08	Meaningful human control	22	0.04

Source: Author (2018)

[59] The author excluded stop words, and excessively vague n-grams.

Most importantly, the results allow estimating states' positions within each of the four issue-areas, and inferring the dependent variable *State preference*. This, in turn, enables establishing the order of each state's preferences and placing them on the described policy continuum (Bueno de Mesquita 2013: 105–109). The table on the following page summarizes the number of observations in each category for each state and scores of *State preferences* (see Table 9). Firstly, the results suggest that discussions have, so far, focused predominantly on international law (50.9%), followed by aspects of operational safety (19.6%) and societal security (16.7%). Curiously, challenges in the area of international security (12.8%) constitute only a small part of the debate, though this is the only category which scores significantly more negatively compared to others. Secondly, the overall score of 0.44 suggests that states would, collectively, prefer hybrid regulation over no regulation or forms of prohibition.[60] Nevertheless, to determine, whether this is a likely outcome of the discussions, the identification of the median voter is necessary. Before so-doing, the thesis further elaborates on the content of states' statements at the multilateral forums.

With regard to the area of international law, statements often acknowledge the difficulties of AWS' compliance with IHL. Nevertheless, the prevalent opinion is that AWS may at some point be able to comply with principles of international law (French Delegation 2015d; German Delegation 2014c; Israeli Delegation 2014). In this regard, states typically argue that the eventual compliance is mostly a matter of gradual technological improvements. Certain delegations, such as Israeli and US delegation, raised the idea that future AWS can actually improve compliance with IHL, whether through the use of self-deactivating and self-destructing munition or by providing greater precision, and thereby reducing the risk of collateral damage or civilian casualties (Israeli Delegation 2015b; US Delegation 2017a). On the other hand, certain other states such as Brazil and Japan questioned the ability of AWS to comply with IHL (Brazilian Delegation 2014; Japanese Delegation 2014). Germany, for example, also said that, under current circumstances, AWS would probably fail

[60] Mathematically, the order of the overall state preference can be expressed as ($|0.50 - 0.44| < |0.25 - 0.44| < |0.75 - 0.44| < |0.00 - 0.44| < |1.00 - 0.44$) based on comparison to the ideal points identified on the policy continuum (see Figure 12).

to pass Article 36 weapon reviews, and would thus be illegal (German Delegation 2015b).

Table 9: Content Analysis Results

Country	International law		International security		Societal security		Operational safety		State preference	
	Count	Score	Count	Score	Count	Score	Count	Score	Count	Score
CHN	5	0.50	4	0.83	5	0.50	2	0.50	16	0.58
USA	47	0.27	17	0.15	5	0.25	24	0.25	93	0.23
IND	24	0.58	18	0.86	9	0.50	8	0.50	59	0.61
RUS	13	0.18	1	0.50	3	0.10	NA	NA	17	0.20
JPN	20	0.55	5	0.50	13	0.15	12	0.42	50	0.41
KOR	14	0.43	5	0.33	7	0.20	2	0.50	28	0.37
BRA	12	0.79	2	0.83	9	0.83	2	0.83	25	0.82
DEU	71	0.23	5	0.75	25	0.82	12	0.43	113	0.56
GBR	26	0.24	4	0.75	3	0.50	9	0.14	42	0.41
FRA	62	0.19	11	0.62	24	0.42	38	0.64	135	0.47
ITA	16	0.46	6	0.50	6	0.36	6	0.50	34	0.46
ISR	28	0.15	7	0.28	2	0.00	15	0.13	52	0.14
Total	338	0.38	85	0.58	111	0.34	130	0.44	664	0.44
Freq.	50.9 %		12.8 %		16.7 %		19.6 %		100 %	

Source: Author (2018)

In addition, several states have expressed their opinion on particular forms of AWS regulation. Crucially, most states of case selection have argued against the case for a preemptive ban, general limitations and all-encompassing regulations (German Delegation 2017b; Russian Delegation 2017; French Delegation 2016a; Israeli Delegation 2015a; Italian Delegation 2016; Korean Delegation 2015b; UK Delegation 2016c). The UK has argued that "[...] it is too soon to ban something we simply cannot define [...] we believe that existing international humanitarian law is sufficient in assessing whether any future weapon system including LAWS would be capable of legal use" (UK Delegation 2016c). Lack of support on the basis of the definitional issues is also emphasized in statements of other states such as Russia (Russian Delegation 2017).[61] Other states argue that preemptive prohibition will "[...] hinder the development of technologies with very useful civilian application" (Italian Delegation 2016). Furthermore, states highlight the fact that the ban would be premature with regard to the fact that the technology still evolves and there are no functional models of fully autonomous systems in existence for now (German Delegation 2017b; Russian Delegation 2017).

As for other solutions to AWS, there seems to be no agreement among states on what alternative to a preemptive ban should be enforced. However, states often do express opinion on the best way to regulate AWS. The US has encouraged other states to adopt national policies similar to DoD's Directive 3000.09 and share best practices on implementation of weapons reviews processes (US Delegation 2015; 2014a; 2014b; 2013). France and Germany suggested that the GGE at CCW should draft a "political declaration" on appropriate transparency and confidence-building measures, including exchange of information on national Article 36 procedures, or, eventually, participation in demonstrations of AWS technology, code of conduct or eventual establishment of a CCW-based permanent committee of technical experts on AWS for further guidance (French Delegation 2017; German Delegation 2017a). On the other hand, the US and UK also argued that IHL and legal weapons reviews on their own provide a sufficient framework for regulation of AWS (UK Delegation 2016a; 2015a; US Delegation 2017a; 2017b). Finally, although Brazil joined CSKR's call at the end of 2017, China has officially adopted the position only in 2018, thus no explicit calls for a ban were captured in the selected timeframe (CSKR 2018b).[62]

[61] The Russian delegation repeatedly invoked this issue as a reason for not discussing the topic of AWS yet, claiming that "[t]he efforts to take course of discussing international legal regulation of using virtual technology that presently has no functioning models seems to be doubtful" (Russian Delegation 2014a). In a similar fashion, Russia warned that discussing the topic prematurely would be unwise due to building expectations as the discussions proceed (Russian Delegation 2014b). On the topic of the Protocol on Blinding Lasers as a model for preventive prohibition on AWS, the Delegation suggested that this can hardly be replicated due to the fact that the Protocol bans a particular employment of the technology—that is, permanent blindness —rather than the technology as such (Russian Delegation 2017).

[62] Several points regarding the notion of MHC should also be mentioned. For most countries of case selection, the notion presents a challenge due to its vagueness, and this could have a negative impact on clarity of discussions (French Delegation 2016c; 2015c). For the US, the vagueness of the vague restriction could actually negatively affect commander's ability to achieve military objectives (US Delegation 2014b). Russia argued that it would be impossible to set a threshold for what would cease to be a meaningful human control (Russian Delegation 2017). For Germany, the notion of

As regards the various challenges to international security identified in the academic literature, the lack of focus on security issues seems to be striking. Even though most states acknowledge that AWS may present some security challenges, explicit judgement on AWS as an actual threat to national or international security is close to absent. To the extent that states mention such issues, the most frequent ones include the potential of proliferation and military effectiveness or utility (see US Delegation 2017a; Chinese Delegation 2016b; French Delegation 2016d; Indian Delegation 2015b; Israeli Delegation 2016; Korean Delegation 2014). However, in eyes of some countries, the US and Israel in particular, AWS provide mostly a way to improve security rather than a challenge (Israeli Delegation 2014; US Delegation 2017a). The US delegation highlighted that the precision of AWS may enhance the ability to reduce casualties, achieve military objectives, and autonomous decision-making and information-processing can provide commanders with time to think about the context (US Delegation 2017a). In a similar fashion, Korean Delegation emphasized use of AWS in dull, dirty, and dangerous missions (Korean Delegation 2014).[63]

On the other hand, the concerns for negative security consequences of AWS were expressed, perhaps surprisingly, by the Chinese Delegation, which described AWS in one of its statements in terms of "[...] immense potential risks and threats to security and the survival of mankind [...]" (Chinese Delegation 2016a). China has further called on the international society to "[...] carry out preventive diplomacy, check the emerging arms race [...] and safeguard international peace and stability" (Chinese Delegation 2016a), and highlighted the undermining of a global strategic balance and stability due to R&D associated with AWS (Chinese Delegation 2016b). For India, the main security concern found throughout most statements is related to the consequences of dissemination, or proliferation, of AWS technologies. India has also repeatedly called for debates to focus greater attention on security issues and threats posed by the technology (Indian Delegation 2016a 2015a; 2014a). In a similar vein, Germany invoked a possible new arms race as a consequence of AWS

---

human control in general presents a fundamental principle of IHL—that is, only humans can decide on life and death decision (German Delegation 2014a). Interestingly, the Indian delegation suggested that adopting the MHC could actually "legitimize" AWS (Indian Delegation 2014b). In regard to MHC, the academic fervor for the notion therefore seems somewhat unwarranted.

[63] This includes, particularly, explosive ordnance disposal or detecting of chemical, biological, radiological and nuclear agents (Korean Delegation 2014).

development and deployment (German Delegation 2015b). Still, however, the lack of focus on issues of international security in countries' statements obfuscates the dynamic of the security dilemma.

With respect to issues of societal security, statements typically concern the usefulness of the underlying technology of AWS in civilian applications—securing the wellbeing of state's society—as well as ethical concerns associated with ceding the decision over life and death to AWS. On the one hand, majority of states seem to agree that discussions on AWS at the forums must not hamper the development of robotic and autonomous technologies for civilian purposes (French Delegation 2014b; Japanese Delegation 2014; Korean Delegation 2015b; 2014; US Delegation 2014c). Japan noted that the underlying technology "[...] is not only useful but also contributes to a safe and sound life for us all" (Japanese Delegation 2014).[64] Korea also suggested that no restrictions shall apply to robotic technologies for peaceful purposes as a result of the debate (Korean Delegation 2014). Despite recognizing a variety of benefits to society, however, most countries seem to agree that the dual-use nature of AWS poses a problem in distinguishing between systems for military and civilian application (French Delegation 2014a; 2014b; Indian Delegation 2017; Korean Delegation 2015a; Russian Delegation 2017). As an extension of this issue, Germany brought up the issue of differentiation between offense-defense purposes (German Delegation 2014a).

On the other hand, the issue of ethical and moral challenges was reflected in various statements, albeit to a lower extent compared to the benefits. For example, the German delegation has argued that maintaining human control over decisions to kill constitutes an indispensable ethical and legal principle (German Delegation 2015a; 2014a; 2014b; 2014d). Brazil mentioned the effects of dehumanization due to the progressive distancing between operators and targets, "risk-free" warfare or issues of accountability for killings and dictates of public conscience as an ethical guidance to AWS use (Brazilian Delegation 2013; 2014). In one of its statements, Brazil brought up a legend about a creature called "Golem", drawing a comparison to ethical wisdom of AWS development (Brazilian Delegation 2014). However, the

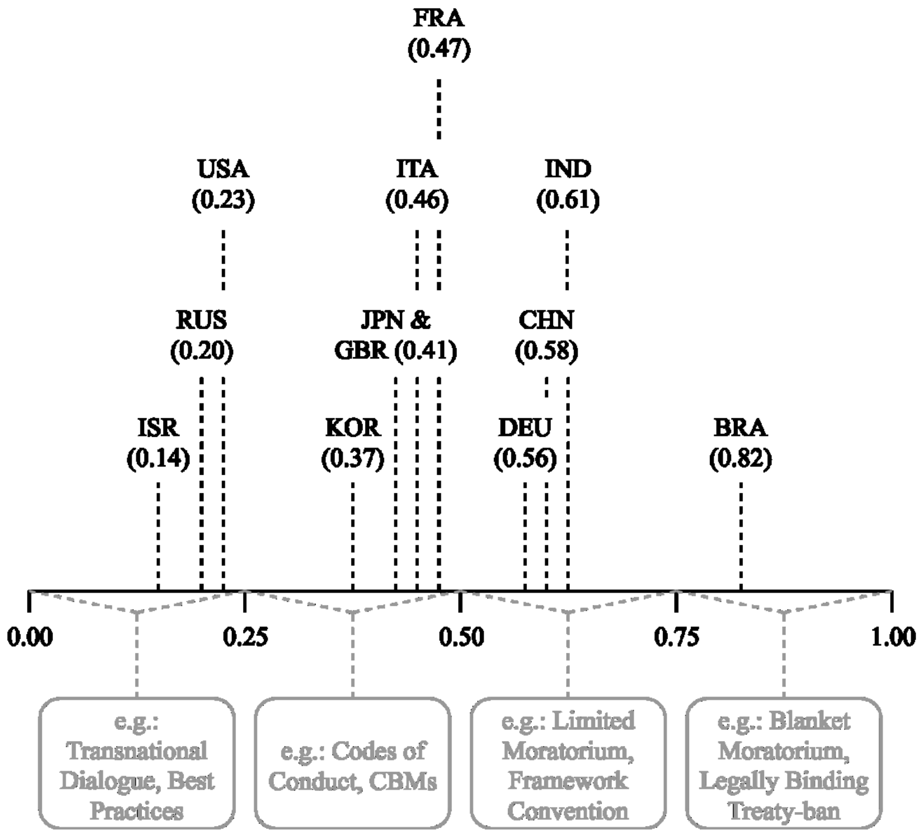
[64] As an example, the Japanese Delegation invoked use of robotic systems in the aftermath of the 2011 Fukushima disaster (Japanese Delegation 2014).

overall attitude to ethical aspects of AWS seems to favor the notion of AWS' ability to actually improve ethical conduct. In this regard, France highlighted the fact that the lack of emotions in autonomous decision-making avoids the pitfalls of killings due to soldiers' vulnerability to negative feelings (French Delegation 2016b; Israeli Delegation 2016). Finally, Russia argued that humanitarian concerns alone do not provide a sufficient basis for banning a weapon (Russian Delegation 2017).

As regards the aspects of operational safety, the discussions have focused mainly on issues related to technical feasibility of AWS' compliance with IHL and the impact of AWS on military effectiveness. To be sure, most views exhibit certain proclivity to technological determinism. States such as the US and Israel suggested that the negative impact of issues of complexity and unpredictability will, at some point, be minimized through technological improvements, for example, by greater precision or effective training of AWS' operators (US Delegation 2014b). Generally, states see the potential in AWS to improve essentially upon every other issue-area, including the improvement of compliance with IHL, ethical conduct and, certainly, military effectiveness also (Israeli Delegation 2015b; 2014; UK Delegation 2016b). That being said, under current circumstances, states perceive high unpredictability as, essentially, rendering the weapon system militarily useless (French Delegation 2016a; 2016c; 2015a; 2015b; 2014c; German Delegation 2014c). It is assumed that no available technology can account for all the environmental complexities, or has adequate sensory capabilities to precisely distinguish between combatants or non-combatants (French Delegation 2016b; German Delegation 2014c).

These divergent views on AWS in the four identified issue-areas show that after more than five years of informal negotiations, there is little consensus among key states on every major issue of AWS R&D. Arguably, this divergence reflects the severity of distributional concerns and various subjective security demands which seem to be further complicated by domestic situations. The distributional concern can be evidenced, for example, by India's repeated statements that the discussions on AWS should be conducted "[...] in a manner that does not widen the technology gap amongst states [...]" (Indian Delegation 2016b; 2015b). In a similar vein, there has been a concern particularly in countries with highly advanced robotics such as Korea and Japan that AWS discussions could lead to restrictions that would negate their relative advantage (Japanese Delegation 2014; Korean Delegation 2015b). As for the subjective security demands, arguably, states such as Israel and the US have shown much greater appreciation for military application of AWS, even though this is hardly surprising with regard to the available strategic documents of the US (see US Army 2017; DoD 2013; 2007), and the "public secret" drone operations by both countries, particularly against terrorists and insurgents (see Ewers et al. 2017).[65]

Figure 14: State Preferences of AWS Regulation[66]



Source: Author (2018)

With states' most-preferred outcomes placed on the continuum, we can see that the landscape translates to roughly 3 to 4 groups with different interests. The first one comprises of the US, Russia and Israel, all of which have shown very little

[65] As an example of domestic pressures, we can, again, recall the example of Japan, which described the underlying technology as a key to ensuring its society's wellbeing (Japanese Delegation 2014).
[66] The exact order of preferences can be found in the appendix (see Appendix 3).

concern over AWS' potential violation of IHL and ethical issues, on the contrary, they seem to agree that AWS could actually improve upon both of these issue-areas as a result of technological development. In addition, the US and Israel have shown interest in military application of AWS. Crucially, all of these states seek greater flexibility, or minimum-obligation outcomes, probably, through minor regulation. They would rather have no regulation over hybrid regulation. The second larger group is much more diverse in terms of opinions on particular subjects. Most are not necessarily concerned with AWS' ability to comply with IHL, and they seem to agree that there are some legitimate security concerns. Where the opinions start to get really different is in the issue-area of societal security and operational safety. While certain states such as Japan and Korea seem to be very enthusiastic about the civilian benefits of the underlying technology, Germany emphasizes the moral dilemmas of autonomous decision making that could lead to death of humans. In addition, the UK sees the issue of operational safety much like the first group. What binds those countries together is a majority preference towards hybrid regulation and preference of minor regulation over prohibition. Most of the countries seem to prefer some degree of lower- to medium-obligation commitments.[67]

The last group comprises of China, India and Brazil. All three states express some concern over AWS' ability to comply with IHL and its ethical and operational implications. Nevertheless, where the group stands out compared to the other two is primarily in the domain of international security. All would seem to prefer some form of partial prohibition of AWS over minor regulation—the preferred outcome of the first group. With the recent joining of China to the CSKR movement for a ban, the prospects of a ban, or partial prohibition, therefore suddenly seem much more realistic as the power on the continuum becomes more evenly distributed. As was previously mentioned, however, estimation of preference must be interpreted with caution because statements reflect what countries want others to know or believe about their policy preferences rather than "actual" position (see Jervis 1988: 323; Schelling 1956: 282). How can we possibly find when a country is telling the truth and when it is bluffing? If offense and defense is indistinguishable as in the case of

[67] Germany and Korea constitute the outliers of the group. In particular, unlike the rest of the group, Korea would seem to prefer minor regulation over hybrid regulation and Germany partial prohibition over minor regulation.

AWS—in fact, offensive capabilities are still to be demonstrated (Altmann & Sauer 2017; Lachow 2017)—states can signal their benign intentions by participating in arms control negotiations, by adopting unilateral defense and restraint (see Glaser 1995). Nevertheless, the current state of discussion on lethal autonomy resembles negotiations about arms control negotiations rather than actual arms control talks.

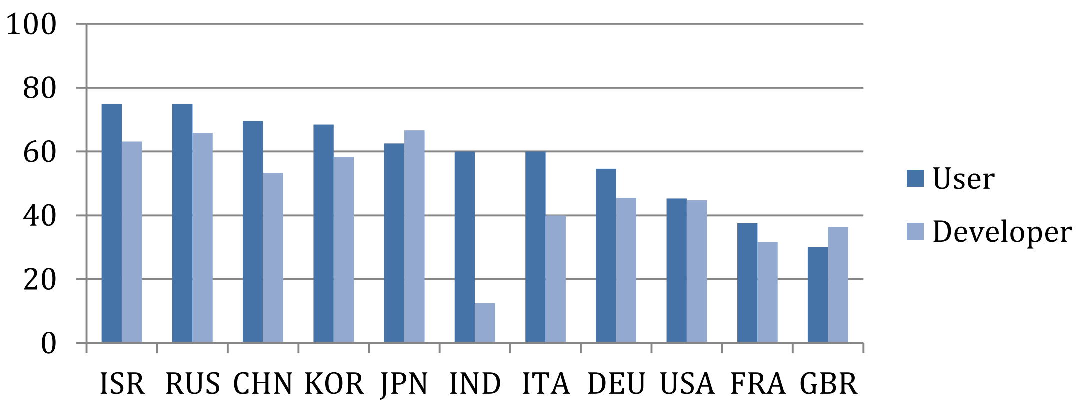
The first way to look at states' signaling benign intentions is to look at what they have said about their own R&D related to AWS. In this respect, multiple states have expressed they do not have an intention to develop and use AWS. Japan noted that "[...] Japan's Ministry of Defence has no plan to develop robots with humans out of the loop, which may be capable of committing murder" (Japanese Delegation 2015). In a similar vein, the UK announced it does not plan to develop any "lethal" AWS while Germany said it would not develop any systems with autonomy over a kill decision (German Delegation 2014c; UK Delegation 2016c; 2015b). Regardless, it might be unwise to take statements like these at their face value. Perhaps a more credible way to show unilateral restraint is to commit oneself to a non-threatening defensive policy such as the ban on AWS. Again, this is the case of Brazil and more recently China as well. Though, a pledge to a ban can, in some sense, be interpreted as a bargaining strategy to gain additional benefits in negotiations than benign self-restraint (Schelling 1956: 288). Furthermore, we can look at countries' attitudes to the process at the CCW, and whether they are supportive of elevation of the debate to a more formal setting akin to arms control negotiations. In this regard, all states ultimately supported the decision to proceed with discussions by establishing the GGE. However, before deciding not to break consensus, Russia did oppose the step (Russian Delegation 2017).[68]

Finally, we can look at countries' actual material capabilities. As mentioned in the Chapter 2, a vast majority of systems with autonomy in certain functions in use today serves only defensive purposes such as the US Phalanx CIWS air defense system (see Scharre & Horowitz 2015). Loitering munitions, or suicide drones, are

[68] Attendees of the 2017 GGE meeting have suggested that Russia is not interested in making any progress on the issue, deliberately attempting to undermine the discussions. In addition, the Russian delegation allegedly said it would not adhere to any international regulation of AWS resulting from the CCW discussions (Tucker 2017). In this light, Russian behavior could, arguably, be characterized as that of a greedy state for which costs of arms control are always perceived as higher compared to pure security-seekers (Glaser 1995: 65–69).

the only example of offensive autonomous systems, and Israel is the single known "user" of these systems (Boulanin & Verbruggen 2017: 50–53; Scharre & Horowitz 2015: 13). Technically speaking, some encapsulated torpedo mines used by Russia and China also constitute human-out-of-the-loop systems which could be used for offensive purposes (Scharre & Horowitz 2015: 14–15). Furthermore, when looking at the sheer number of military robotic or autonomous systems at states' disposal, this raises questions of credibility of the signals of benign intentions. According to available data on autonomous systems, we can see that a good portion of systems is weaponized. This is particularly the case of Israel, Russia, China, Korea, Japan or Italy (see Figure 15). Obviously, a large part of systems are dual-use and offensive and defensive purposes are not clearly distinguishable but the credibility of states' signals seems doubtful at the backdrop of the pursuit of R&D of highly autonomous weaponized systems.[69]

Figure 15: Proportion of Armed Systems with Autonomous Functions (%)[70]



Source: Author (data: Boulanin & Verbruggen 2017)

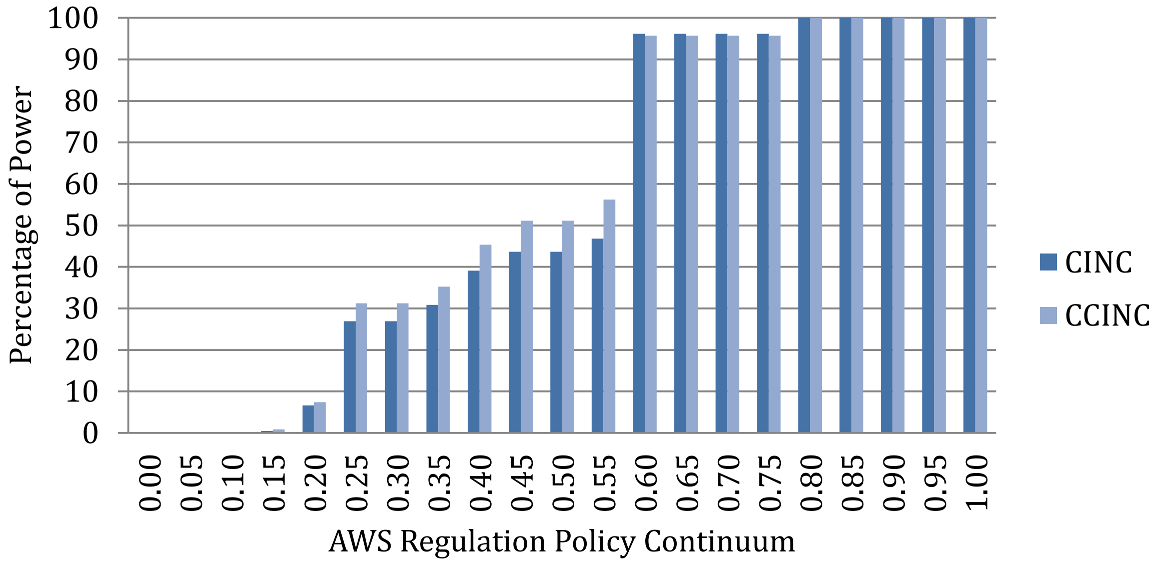
[69] While certain defensive realists such as Shiping Tang would criticize the fact that an objective offense-defense balance is a hoax, they nevertheless maintain that perceptions of offensive-defensive postures matter (Tang 2010: 222, 240). By the same token, when an adversary arms its systems, any rational actor would arguably question whether the system is genuinely defensive. This might not necessarily present an issue in case of static AWS such as robotic sentries or active protection systems but most systems with autonomous functions such as UAVs, UGVs and UUVs exhibit high mobility (see Boulanin & Verbruggen 2017; Scharre & Horowitz 2015).

[70] The author of this thesis would like to thank SIPRI staff for providing the dataset on autonomous systems. The dataset at hand was originally intended to accompany the release of the publication by Boulanin and Verbruggen (2017) (see Boulanin & Verbruggen 2017). Data cover arsenals of robotic and autonomous weapon systems of mostly 12 states, including the ones displayed.

Regardless of states' fundamental intentions, the most likely outcome of the discussions will depend, to a large extent, on the power behind particular solutions to AWS regulation. The thesis thus employs the MVT model as described by Bueno de Mesquita (2013). To find the median voter, or the winning outcome, we can use a simple graph of cumulative total amount of power, or the votes, plotted against a policy continuum of AWS regulation. Further, the policy continuum is divided into fifth of a point swaths to capture particular policy positions. The median voter can then be found at the point where the amount of cumulative total power reaches the 50% mark since this is the position in the middle (Bueno de Mesquita 2013: 111–112). Operating under the realist assumptions of primacy of hard power, the thesis first turns to a model, which uses CINC and CCINC scores (see Figure 16). As shown on the graph on the following page the outcome differs with use of different types of power. Particularly, CINC predicts the policy position at 0.60 as the outcome, whereas CCINC predicts the position at 0.45 as the outcome. Put differently, were the results of the current debate dependent purely on material factors—that is, on the use of threats and payments (see Nye 2011)—the most likely outcome of the discussions would be hybrid regulation, representing a near-ideal point of India, China and Germany.

However, the results turn differently when the role of context is introduced into the model. If we account for the role of AWSIB, which represents states' stakes in the discussions, issue-salience and the ability to commit credibly to a particular policy position (see Lebow 1996; Schelling 1956), the most likely outcome would be, again, hybrid regulation, though the particular form would reflect most closely a near-ideal point of the UK, France, Italy and Japan. Interestingly, such an outcome would reflect interests of the least powerful of the three groups—that is, at least in terms of the total share of power across the continuum (see Appendix 4). Although, it is perhaps not unreasonable to suggest that this is actually, in certain sense, also the most powerful group of the three. First, the appearance of "weakness" can give the group greater bargaining power in resolving distributional problems linked to potential regulatory solutions, for example, by appearing unable to back down due to domestic pressures. Second, states' choice of the policy position in the "middle" may actually be a conscious and deliberate choice, or strategy, as it represents the most secure position (Bueno de Mesquita 2013: 112–113; Schelling 1956: 288). To provide an alternative perspective, we can use the earlier identified "smart" power measures of SPI and CSPI to see how the results differ under assumption that soft power factors of persuasion and attraction (see Nye 2011) play a role in the likely development of the discussions on AWS. The graph on the following page predicts the position at 0.40 as the winning outcome in employment of both measures. This is a position most closely linked to ideal points of the UK, Japan and Korea, which, again, represent the "weakest" power bloc in case selection.

Figure 16: Median Voter Prediction based on Cumulative Hard Power[71]



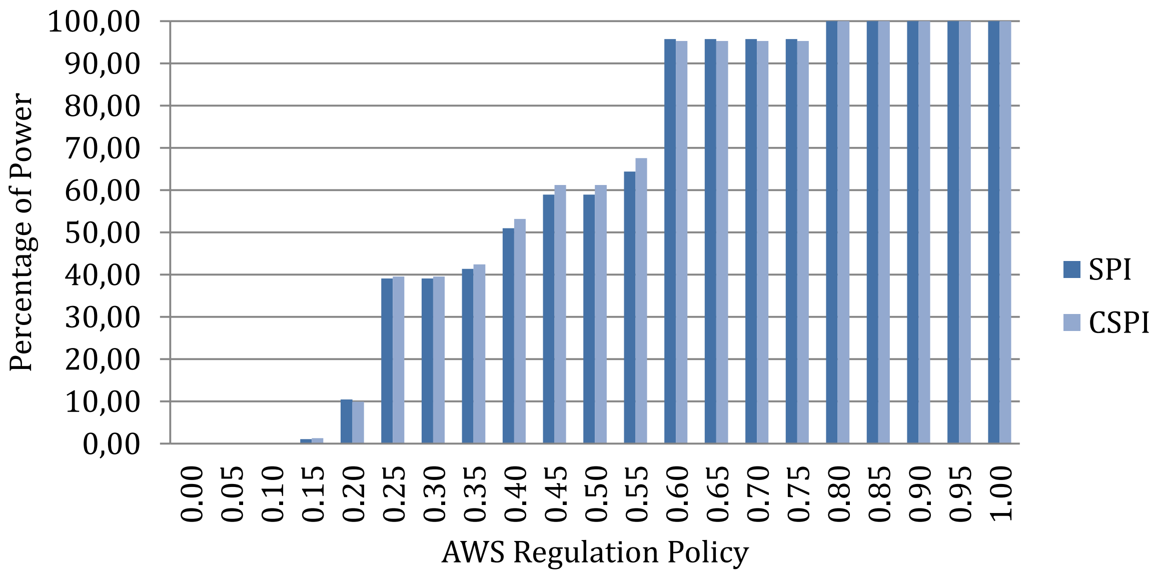
Source: Author (2018)

With these results, we can confirm the first two hypotheses and answer the first research question. First and foremost, all four models of MVT prediction show that countries would prefer some form of regulation over no regulation at all, both individually and collectively. To be sure, certain states, such as Russia, Israel or the US, have shown very low support to more restrictive forms of regulation, let alone

[71] It should be noted that the graph operates with shares of power within the boundaries of case selection. In other words, the results would be most precise if only the countries of case selection would have the power to vote. However, the model represents only 62.7% and 69.2% of world totals respectively for CINC and CCINC. Use of the model is justified by the fact that it includes only the most powerful countries in the context at hand.

prohibition of AWS but none have explicitly rejected the potential regulation in the official statements. Second, we confirm the second hypothesis. The most powerful state as regards the indicators of hard power, China, would overall secure its ideal policy position, which is, at least during the selected timeframe, hybrid regulation. However, perhaps unsurprisingly, the precise position is closer to the ideal point of the "weakest" group of states which place themselves in the middle. Furthermore, if we were to adopt the alternative model using measure of smart power, we would have to reject the second hypothesis because, even though the outcome of hybrid regulation remains unchanged, the power distribution shifts in favor of the US that would be unable to secure its most-preferred outcome—minor regulation. In both approaches, however, the "weakest" power bloc occupies the precise median voter position, and in both approaches, the role of the context affects the results.

Figure 17: Median Voter Prediction based on Cumulative Smart Power



Source: Author (2018)

Which of the proposed regulatory solutions—ranging from no regulation to total prohibition—is likely to garner the most support? Irrespective of a particular measure used for median voter outcome prediction, this analysis suggests that the most likely outcome is some form of hybrid regulation. On average, states seem to prefer lower- to medium-obligation forms of AWS regulation, combining elements of both soft and hard law. This can include, on one end, highly formalized forms of transparency and confidence-building measures or codes of conduct and, on other end, moratoria limited to particular timeframes or specific uses of AWS and stages of development, acquisition, R&D, testing, deployment, transfer, and so on. Further solutions might also include creating a framework convention to enable countries to adopt additional protocols on a voluntary basis (see Crootof 2016a: 1899; Jenks 2016: 59–62; Geiss 2015: 24; Heyns 2013: 21; Marchant et al. 2011: 313–314). It is evident from the substantive content of countries' statements in debates at various forums that flexibility is more than desirable. Hybrid regulation presents a modest form of regulation that should not impede the technological progress, retaining the right to revise any restrictions in the face of legitimate, security-seeking interest of involved parties, when circumstances change. At the same time, it is a compromise between positions of the two most powerful states, the US and China, which can be viewed as a necessary precondition for agreement's effectiveness. While explaining states' preferences is out of scope of this thesis, especially due to constraints of the sample size, we can at least explore drives of states' interest in the discussion. This is the purpose of the following section.

6.2. Drivers of Discussions on AWS: States' Institutional Activity

Before elaborating on the results of the regression analysis, it is important to make several clarifications. First, the thesis provides 4 models of OLS regression. These include Model 1, Model 2 and respective modifications with robust standard errors (RSE). The crucial difference between these two models lies in the inclusion of the variables *R&D spending* in Model 1 and *Military expenditure* in Model 2. The reason behind separation of the two variables is that both are highly correlated, at 0.9 level, which could result in severe multicollinearity (Allison 1999: 63). Thus, to test the hypotheses without compromising on the effectiveness of the regression estimation, the author provides two models.[72] Second, the thesis uses logarithmic transformations for variables with highly skewed distributions (Lane 2018). These are, in particular, *GDP per capita*, *Institutional activity* and *Military expenditure*. The author also provides two modified models using the RSE to preclude any potential heteroscedasticity issues (Monogan 2015: 89). All independent variables used in regression are averages for the whole period from 2013 to 2017 when possible. However, due issues with data availability, *Militarization*, *R&D spending* and *UCAV program* cover only the years 2013–2016, 2013–2015, and 2013–2014.

Table 10: Regression Analysis Results

	Dependent variable: Institutional Activity			
	Model 1	*Model 1 (RSE)*	*Model 2*	*Model 2 (RSE)*
Conflict	0.028	0.028	0.051	0.051
	(0.071)	(0.066)	(0.048)	(0.046)
GDP per capita (log)	-0.020	-0.020	0.168***	0.168**
	(0.110)	(0.137)	(0.062)	(0.068)
UCAV program	0.528***	0.528***	0.543***	0.543***
	(0.193)	(0.190)	(0.174)	(0.196)
Democracy	0.049***	0.049**	0.024**	0.024*
	(0.016)	(0.021)	(0.010)	(0.014)
Ban	0.694***	0.694***	0.561***	0.561***
	(0.205)	(0.181)	(0.148)	(0.162)
R&D spending (log)	0.199***	0.199***		
	(0.045)	(0.049)		
Military expenditure (log)			0.181***	0.181***
			(0.035)	(0.038)
Militarization	-0.013	-0.013	-0.028**	-0.028**
	(0.016)	(0.016)	(0.012)	(0.013)
LDCs	0.441	0.441	0.304*	0.304*
	(0.332)	(0.354)	(0.168)	(0.165)
Constant	-1.949***	-1.949**	-3.543***	-3.543***
	(0.725)	(0.813)	(0.506)	(0.495)
Observations	80		137	
R^2	0.639		0.626	
Adjusted R^2	0.598		0.602	
Residual Std. Error	0.618 (df = 71)		0.605 (df = 128)	
Note:				*p<0.1; **p<0.05; ***p<0.01

[72] For correlation matrix please see Appendix 5.

Based on the results of Model 1, we can confirm hypotheses 5, 7 and 9. That is, there is significant positive relationship between *Institutional activity* and *UCAV program*, *Democracy* and *R&D spending*. The variable *Ban* proved highly significant which suggests that supporters of a ban belong among the more active participants of the discussions. Based on the results of Model 2, we can confirm hypotheses 4, 6, 7 and 9. Thus, the model suggests that *Institutional activity* is positively correlated with *GDP per capita*, *Military expenditure*, *UCAV program*, and *Democracy*. The *Ban* variable again proves to be highly significant and even the control variable *LDCs* is somewhat significant in this model. Furthermore, we must reject hypotheses 3 and 8 since the variable *Conflict* proved insignificant in both models and the variable *Militarization* exhibits "negative" rather than positive correlation with *Institutional activity* in the second model. With regard to models' explanatory value, we can say that Model 1 explains 63.9% of variance in *Institutional activity* whereas Model 2 62.6%. However, if we account for unexplained independent variables, adjusted R^2 suggests that Model 2 is actually the better of the two, with 60.2% when compared to 59.8% of Model 1. The results remain almost unchanged when using RSE which increases credibility of the findings.[73]

In summary, the regression analysis finds that democratic states with high GDP per capita, R&D and defense spending with programs of armed drones are the most active participants in the discussions on AWS and, by extension, also exhibit the greatest "interest" in subject in question. The significance of the variable *UCAV program* constitutes perhaps the most interesting finding in this regard. It is worth mentioning that the variable in bivariate linear regression model with *Institutional activity* explains over 33% of variance, and its Pearson correlation score is at 0.54 level—by far the highest correlation from all independent variables (see Appendix 5). On one hand, the finding is hardly surprising given that states with programs of UCAVs are likely the most invested in AWS-related R&D and thus have the greatest stake in discussions because the distributional consequences of the potential ban would likely affect them disproportionately (Hansel et al. 2018: 6). Nonetheless, on the other hand, if UCAVs are understood as precursors to future AWS, this begs the

[73] In addition, it should be noted that the results of Model 1, and, particularly its inability to explain significance of certain variables, is probably due to the lower number of observations and because of multicollinearity issues stemming from inclusion of the variable R&D spending.

question, what do the countries with such programs seek to achieve in the debates at HRC, UNGA, and CCW? Are they genuine security-seekers or greedy states which have no intention to abide by the potential agreement?[74]

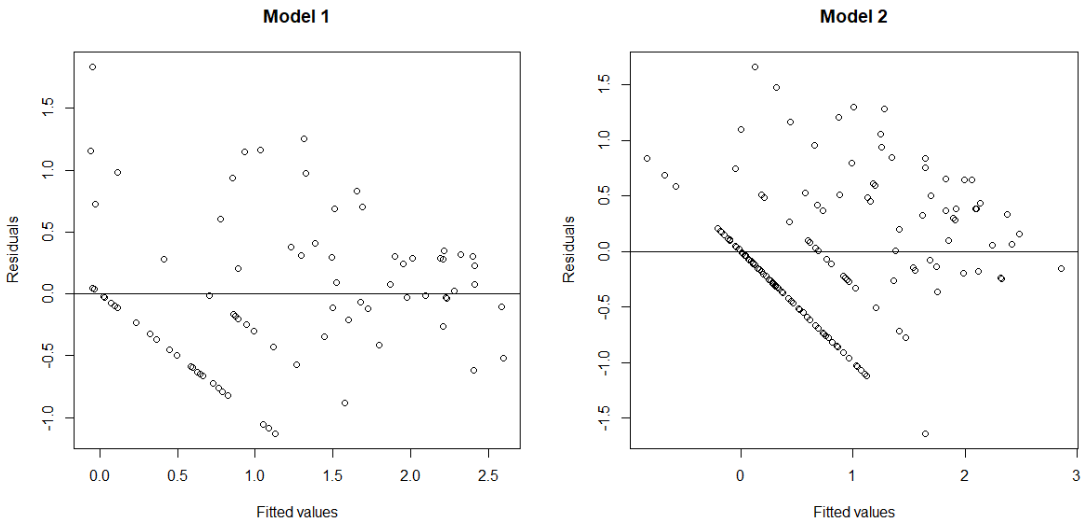
As regards others variables, first, we can speculate that the variable *Conflict* did not prove to be significant because, although states facing external threats may have incentives to build arms, engagement in arms control discussions—or in our case informal discussions at the fora for arms control—would likely constrain their military options and undermine the main interest of survival (Mearsheimer 1995: 11–12; Walt 1998: 31; Waltz 1979: 126). Indeed, Bernauer & Ruloff (1999) suggest that countries facing external threats are unlikely to engage in negotiation of arms control agreement (Bernauer & Ruloff 1999: 11–16). With regard to *Militarization*, and its negative correlation with *Institutional activity* in the second model, we can suggest that this approximation of states' ambitions to project power also captures their greediness and thus incentives to pursue non-cooperative aggressive policies rather than defensive ones such as unilateral restraint or arms control (see Glaser 1997; 1995). In other words, subjective security demands of the most militarized states exacerbate the distributional concerns to the point where even participation in informal negotiations might dissuade them from cooperating (Glaser 1995: 65–69).

While the effects of *Democracy*, *R&D spending* and *Military expenditure* are relatively straightforward, the two control variables *Ban* and *LDCs* provide a more interesting subject of exploration. First, with regard to the control variable *LDCs*, it could be argued that the distributional consequences operate similarly as with the most advanced countries. The concern for a "widening technological gap" was also emphasized in the previous section in one of India's statements (Indian Delegation 2016b; 2015b), though no LDC was included in the case selection for the content analysis. As for the variable *Ban*, it is questionable whether the resulting effect can be attributed to political-ideological activism or rather a pronounced distributional concern. In any case, it should be mentioned that in a bivariate linear regression, it

[74] Of the 25 countries with UCAV programs in the dataset by Fuhrmann and Horowitz (2017b), only two support a ban as of today: China and Pakistan. China's position on AWS regulation presents an intriguing puzzle in this regard. If we accept Glaser's hypothesis about signaling benign intentions when offense and defense are indistinguishable, China's call for prohibition could be interpreted as a genuine signal of benign intentions to ameliorate the security dilemma (Glaser 1995: 65–69).

only explains 2.2% of variance in *Institutional activity* and its Pearson correlation score achieves 0.2 (see Appendix 5). Compared to dummy variable *UCAV program*, it therefore performs relatively poorly. Finally, it should be mentioned that other explanations exist for the significant positive effect of Democracy other than those described in the theoretical framework (see Chapter 4). Democracies are typically more transparent, and allow for free flow of information (see Hollyer et al. 2011). Thus, though the hypothesis here holds, the author is unable to determine whether it is precisely due to greater transparency and openness or pressures due to issues associated with greater public sensitivity to casualties, which drives states interest in AWS-related R&D. Following are the results of several regression diagnostics.

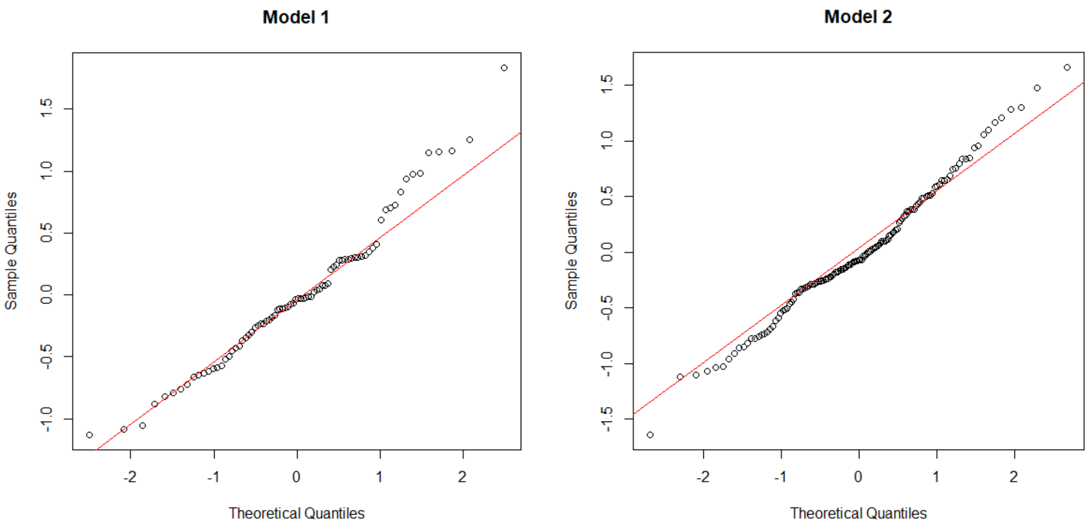
Figure 18: Regression Residuals



Source: Author (2018)

First, the graphs with residuals plotted against fitted values show that there is a clear pattern of rising and falling in both models. As an attempt to find a better functional form, the author tried employing the generalized least squares multiple regression mode but the results did not yield an improvement with regard to other diagnostics.[75] Next, to test whether the models are affected by heteroscedasticity, the author employed the Breusch-Pagan test. With scores of 0.151 for Model 1 and 0.1924 for Model 2 both above the 0.05 threshold, we can reject heteroscedasticity and assume that there is constant variance. Furthermore, to test the assumption of normality, the thesis provides graphs of quantile-quantile plots for the two models (see Figure 15). The array of observations loosely copies the line which shows that residuals are somewhat normally distributed. This is further confirmed by Jarque-Bera test with scores 0.09578 for Model 1 and 0.4384 for Model 2 above the 0.05 threshold, although the first model performs relatively poorly. Lastly, with regard to multicollinearity, the thesis provides VIF scores for both models (see Table 11). Judging from the very low values, we can conclude that there are no severe issues of multicollinearity in both models. However, the relatively higher scores for *GDP per capita* and *R&D spending* suggest that the significance of *GDP per capita* in the first model might have been suppressed by the latter variable. In summary, Model 2 performs relatively better than Model 1 and satisfies the assumptions.

Figure 19: Quantile-Quantile Plots



Source: Author (2018)

[75] Multiple variations of the generalized least squares model have failed the essential test of plotting residuals against fitted values with all observations clustered in the very corner of the scatterplot. In addition, the model failed both Breusch-Pagan test and Jarque-Bera, failing assumptions of normality and homoscedasticity.

Table 11: VIF Scores

Statistic	Model 1	Model 2
Conflict	1.561126	1.277246
GDP per capita	4.076953	3.331263
UCAV program	1.460502	1.644545
Polity	1.374053	1.274541
Ban	1.203197	1.104355
R&D spending	3.404064	NA
Military expenditure	NA	2.595281
Militarization	1.227345	1.154989
LDCs	1.598751	2.018615

Source: Author (2018)

7. Conclusion

What are the prospects of a preemptive ban on AWS? The research suggests that much of the enthusiasm associated with initiatives such as the CSKR has been somewhat unwarranted. More specifically, the likelihood of concluding any legally-binding regulation, let alone preemptive ban, seems negligible. While this research suggests hybrid regulation, including solutions such as moratorium or framework convention, as a more likely outcome, prospects of arms control remain uncertain, nonetheless. The building momentum of the movement for a ban has a potential of exerting pressure on those unconvinced of the need for prohibition but it also has a potential for exacerbating the distributional conflict among participants in debates on AWS due to relative differences in costs and benefits of any such agreement. As demonstrated through the content analysis, which provided an extensive empirical account of the discussions, a consensus is yet to arise among states on how best to regulate AWS. Nonetheless, the overwhelming opinion seems to be that preventive prohibition would be premature, and could hamper the development of associated technologies for civilian application. Although countries acknowledge the various challenges, including legal, ethical and operational aspects of AWS R&D, the lack of focus on issues of international security seems striking.

As for the specific research questions and hypotheses, the thesis was able to provide an answer to the first research question by employing a model for median voter prediction, using different measures of power. When using the "hard" power measure of CINC and adjusted "contextual hard power" of CCINC, the model shows that a policy position of hybrid regulation is the most likely outcome. This allowed the author to confirm the first two hypotheses as the model predicts a policy other than no regulation, and since hybrid regulation reflects the preference of the most powerful state in case selection, which is China. Interestingly, the precise position also reflects interests of the "weakest" group of countries. The author suggests that this is a deliberate strategic choice of by those states, as the position in the middle is the most secure position. Furthermore, the regression analysis provided answer to the second research question. States' activity at the HRC, UNGA and CCW, which is here understood as an approximation of states' interest in the topic of AWS, can be explained by various capacities or incentives associated with AWS-related R&D. This includes enablers such as R&D and defense spending, economic development but also existence of a program on armed drones, which lead the author to confirm four additional hypotheses. The fact that UCAV program has the highest individual explanatory value of states' activity certainly raises some questions as regards the genuineness of states' signaling of benign motives. With these findings, the author concludes that the aims and purposes of this thesis were achieved.

As regards the future research on the topic of AWS or AWS regulation, there are multiple avenues for improvement and building upon the findings presented in this thesis. First and foremost, the results could be improved by further collection of data on states' preferences. While the thesis covered the most important players in AWS discussions, there are still dozens of states which may have an impact over the future regulation, especially if we consider the role they could play in swaying the overall trajectory of the debate, potentially, either towards softer regulation, or more restrictive prohibition. Larger sample size could make the median prediction more accurate and credible and the same applies to estimation of state preferences through the content analysis. Second, the tailor-made framework for assessment of states' preferences presented in the thesis should be tested empirically by looking at states which are likely to join the CSKR call for a ban. Were this framework able to successfully predict which states—based on the higher scores on the regulatory policy continuum—will join the movement, this would improve the robustness of the research design, and author's methodology would increase in credibility. At the very least, other methods of estimation could be adopted to test the reliability, for example expert surveys or interviews with participants of the meetings.

Third, the role of UCAV program as a factor in states' activity at the various international forums deserves greater attention. Although the mere existence of a program does not necessarily rule out legitimate security interests of those states that intend to develop and use these precursors to AWS, as was earlier mentioned, the effect on countries' activity at the forums raises some questions about genuine intentions as regards the desired outcome of the discussions. Are the supporters of prohibition which simultaneously have UCAV programs genuine security-seekers, or are they bluffing with regard to later compliance with the potential agreement? Furthermore, more robust results of the regression analysis could be achieved by extending the coverage of countries and years, and by creating a panel study rather than a simple model using averages for the whole period. Fourth, the role of issue-linkage—as suggested in the original project of the thesis—should be examined. In this regard, more credible estimation of the most likely outcome of the discussions could be made. While the MVT prediction constitutes a useful first step, in practice, arms control typically involves multiple negotiations taking place at the same time over multiple issues, and use of positive and negative inducements by which states can gain concessions from others on issues seemingly irrelevant to the agreement at hand (see Poast 2012; Bernauer & Ruloff 1999; Putnam 1988; Sebenius 1983).[76]

Without a doubt, the issue of "killer robots" presents a fascinating subject of research which deserves greater attention in the field of IR. Throughout the history of mankind, technology proved to have immense transformative effects on society as well as the "international society". If we concede to the idea of inevitably of the technological progress and "homo homini lupus", there is no reason for attempting to stop the "killer robots". To combat the discursive tendencies of "roboutiopia" or "robocalypse", an informed, open and objective academic discussion is necessary, irrespective of a specific theoretical approach. Fully autonomous weapons systems might come tomorrow, in several years or decades—or they may never come at all, yet have already affected the international society in many ways. Whether we look at the lively discussion at fora such as HRC, UNGA and CCW, or rapid investment in robotic and autonomous systems, the growing significance of AWS is obvious. For this reason, further research is necessary to preclude any of the variety of ethical, legal, security and technical issues, ensuring that the underlying technology serves the public interest.

[76] For illustration, while Russia seems currently disinterested in any form of regulation on AWS, other states, including China or the US could bring Russia to joint settlement through side payments such as development or security aid or through sanctions on its high-technology exports. In another scenario, states could simultaneously negotiate the AWS regulation and a trade deal concerning high-technology exports, which would induce the most disinterested states in AWS regulation to re-evaluate the costs and benefits of distributional consequences.

Shrnutí

Autonomní zbraňové systémy v současnosti v oboru Mezinárodní vztahy a Bezpečnostní studia představují spíše okrajové téma. Tato práce představuje jeden z prvních pokusů o systematickou analýzu substantivního obsahu multilaterálních diskuzí na mezinárodních fórech jako Rada pro lidská práva, První výbor Valného shromáždění či Úmluva o některých konvenčních zbraních na půdě OSN. Jedná se současně o první pokus o predikci nejpravděpodobnějšího výsledku těchto diskuzí s ohledem na teoretické poznatky teorie defenzivního realismu, kontroly zbrojení a některých dalších přístupů. Základním předpokladem je, že potenciální důsledky vývoje a nasazení autonomních zbraní vyvolávají bezpečnostní dilema, které vede státy k vyjednávání o kontrole zbrojení. Hypotézu, že státy budou preferovat jistou formu regulace před žádnou regulací těchto systémů, práce dále testuje s využitím metod obsahové analýzy a teorému středového voliče. Výsledky práce ukazují, že nejpravděpodobnějším výsledkem bude mírně závazná forma regulace, „hybridní regulace“, která odpovídá specifickým řešením jako rámcové úmluvy a moratoria. Dalším zjištěním je, že aktivita států na výše zmíněných multilaterálních fórech, která zde představuje „zájem“ o dané téma, souvisí s motivy a kapacitami k vývoji technologií souvisejících s autonomními zbraňovými systémy.

Summary

To this day, autonomous weapons systems have been relatively peripheral subject of interest in the field of International Relations and Security Studies. This thesis presents one of the first attempts at systematic analysis of the content of the debates at multilateral fora, such as the Human Rights Council, First Committee of the General Assembly, or Convention on Certain Conventional Weapons at the UN. The thesis also presents a first attempt to predict the most likely outcome of those discussions as regards theoretical contributions of defensive realism, arms control and certain other approaches. The essential assumption is that potential impact of development and use of autonomous weapons systems cause the security dilemma that leads states to negotiations of arms control. The hypothesis that states would prefer some regulation over no regulation of those systems is tested by employing methods of content analysis and median voter theorem. Results of the thesis show that the most likely outcome is a moderate-obligation form of regulation, “hybrid regulation”, which includes solutions such as framework conventions or moratoria. Additionally, the thesis finds that states’ activity at the forums, which represents their “interest” in the topic, is correlated with incentives and capacities to develop and use technologies related to autonomous weapons systems.

8. References

8.1. Monographies and Chapters

Allison, P. (1999). *Multiple Regression: A Primer* (London: Pine Forge Press).

Altmann, J. – Liebert, W. – Neuneck, G. – Scheffran, J. (1998). Preventive Arms Control as a Prerequisite for Conversion of Military-Related R&D. In: Reppy, J. – Rotblat, J. – Holdren, J. – Avduyevsky, V. eds., *Conversion of Military R&D* (London: MacMillan Press), pp. 255–272.

Anderson, K. – Waxman, M. (2017). Debating Autonomous Weapon Systems, Their Ethics, And Their Regulation Under International Law. In: Brownsword, R. – Scotford, E. – Yeung, K. eds., *The Oxford Handbook of Law*, Regulation, and Technology (Oxford: Oxford University Press), pp. 1097–1117.

Arnold, R. (2017). Legal Challenges Posed by LAWS: Criminal Liability for Breaches of IHL by (the Use of) LAWS. In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 9–18.

Benoit, K. – Laver, M. (2006). *Party Policy in Modern Democracies* (New York: Routledge).

Bernauer, T. – Ruloff, D. (1999). Introduction and Analytical Framework. In: Bernauer, T. – Ruloff, D. eds., *The Politics of Positive Incentives in Arms Control* (Columbia: University of South Carolina Press), pp. 1–46.

Black, D. (1958). *The Theory of Committee and Elections* (Cambridge: Cambridge University Press).

Bongard, D. – Sayers, T. (2002). Impact of Robotics on Nontraditional Warfare. In: Schilling, W. ed., *Nontraditional Warfare: Twenty-First Century Threats and Responses* (Washington, D.C.: Brassey's), pp. 299–309.

Boulanin, V. – Vergbruggen, M. (2017). *Mapping the Development of Autonomy in Weapon Systems* (Stockholm: SIPRI).

Bueno de Mesquita, B. (2013). *Principles of International Politics, 5th Edition* (Washington, DC: CQ Press).

Buzan, B. – Herring, E. (1998). *The Arms Dynamic in World Politics* (Boulder: Lynne-Rienner).

Crootof, R. (2015). The Varied Law of Autonomous Weapon Systems. In: Williams, A. – Scharre, P. eds., *Autonomous Systems: Issues for Defence Policymakers* (Norfolk: NATO Capability Engineering and Innovation Division), pp. 98–126.

Dahl, R. (1961). *Who governs? Democracy and power in an American city* (New Haven: Yale University Press).

Dhanapala, J. (2017). The Security Impact of Lethal Autonomous Weapons Systems. In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 56–65.

Dickinson, W. (2010). Visual displays for mixed methods findings. In: Tashakkori, A. – Teddlie, C. eds., *Handbook of mixed methods in social & behavioral research* (Thousand Oaks: SAGE Publications), pp. 469–504.

Dowding, K. (2011). *Encyclopedia of Power* (Thousand Oaks: SAGE Publications).

Downs, A. (1957). *An Economic Theory of Democracy* (New York: Harper and Row).

Ford, C. (2017). International Humanitarian Law, Article 36, and Autonomous Weapons Systems. In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 76–84.

Hagström, M. (2017). Complex Critical Systems: Can LAWS Be Fielded? In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 128-134.

Horowitz, M. (2017). Military Robotics, Autonomous Systems, and the Future of Military Effectiveness. In: Reiter, D. ed., *The Sword's Other Edge: Trade-offs in the Pursuit of Military Effectiveness* (Cambridge: Cambridge University Press), pp. 161–196.

Keeley, T. (2015). Auditable Policies for Autonomous Systems (Decisional Forensics). In: Williams, A. - Scharre, P. eds., *Autonomous Systems: Issues for Defence Policymakers* (The Hague: NATO Communications and Information Agency), pp. 196–222.

Koh, C. (2017). LAWS in the Maritime Domain: An Asia-Pacific Scenario. LAWS in the Maritime Domain: An Asia-Pacific Scenario. In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 201–216.

Krippendorff, K. (2013). *Content Analysis: An Introduction to Its Methodology* (Thousand Oaks: SAGE Publications).

Krishnan, A. (2009). *Killer Robots: Legality and Ethicality of Autonomous Weapons* (Burlington: Ashgate).

Larsen, J. (2002). An Introduction to Arms Control. In: Larsen, J. ed., *Arms Control: Cooperative Security in a Changing Environment* (Boulder, CO: Lynne Rienner Publishers), pp. 1–15.

Laver, M. (2003). Why should we estimate the policy positions of political actors? In: Laver, M. ed., *Estimating the Policy Position of Political Actors* (New York: Routledge), pp. 3–9.

Lobell, S. (2010). Structural Realism/Offensive and Defensive Realism. In: Denemark, R. ed. *The International Studies Encyclopedia* (Chichester: Blackwell Publishing), pp. 6651–6669.

Mayer, C. (2015). Developing Autonomous Systems in an Ethical Manner. In: Williams, A. - Scharre, P. eds., *Autonomous Systems: Issues for Defence Policymakers* (The Hague: NATO Communications and Information Agency), pp. 65–82.

Mearsheimer, J. (2001). *The Tragedy of Great Power Politics* (New York: W. W. Norton).

Monogan, J. (2015). *Political Analysis Using R* (London: Springer).

Morrow, J. (1994b). *Game Theory for Political Scientists* (Princeton: Princeton University Press).

Mutschler, M. (2013). *Arms Control in Space: Exploring Conditions for Preventive Arms Control* (Basingstoke: Palgrave Macmillan).

Neuendorf, K. (2002). *The Content Analysis Guidebook* (Thousand Oaks: SAGE Publications).

Neuneck, G. and Mutz, R. (2000). *Vorbeugende Rüstungskontrolle: Ziele und Aufgaben unter besonderer Berücksichtigung verfahrensmäßiger und institutioneller Umsetzung im Rahmen internationaler Rüstungsregime* (Baden-Baden: Nomos).

Nye, J. (1990). *Bound to lead: the changing nature of American power* (New York: Basic Books).

Richardson, L. (1960). *Arms and insecurity: A mathematical study of the causes and origins of war* (Chicago: Boxwood).

Rickli, J. (2017). The Impact of Autonomous Weapons Systems on International Security and Strategic Stability. In: Ladetto, Q. ed., *Defence Future Technologies: What we see on the horizon* (Thun: Armasuisse), pp. 61–63.

Sagan, S. (2013). *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons* (Princeton: Princeton University Press).

Schelling, T. (1980). *The Strategy of Conflict* (Cambridge: Harvard University Press).

Sebenius, J. (1983). Negotiation Arithmetic: Adding and Subtracting Issues and Parties. *International Organization* 37 (2), pp. 281–316.

Selim, G. (2013). *Arms Control and Security Cooperation: Contending Approaches* (Berlin: Springer).

Sharkey, N. (2012b). Killing Made Easy: From Joystick to Politics. In: Lin, P. – Abney, K. – Bekey, G. eds., *Robot Ethics: The Ethical and Social Implications of Robotics* (Cambridge: MIT Press), pp. 111–128.

Singer, P. (2009). *Wired for War: The Robotics Revolution and Conflict in the 21st Century* (New York: The Penguin Press).

Snyder, J. (1991). *Myths of Empire: Domestic Politics and International Ambition* (Ithaca: Cornell University Press).

Thee, M. (1986). *Military technology, military strategy, and the arms race* (New York: Free Press).

Wallach, W. (2017a). Predictability and Lethal Autonomous Weapons Systems (LAWS). In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 295–312.

Waltz, K. (1979). *Theory of International Politics* (New York: Random House).

8.2. Peer-reviewed Articles

Abbott, K. – Keohane, R. – Moravcsik, A. – Slaughter, A. – Snidal, D. (2000). *The Concept of Legalization*. International Organization 54 (3), pp. 401–419.

Abbott, K. – Snidal, D. (2000). Hard and Soft Law in International Governance. *International Organization* 54 (3), pp. 421–456.

Adams, T. (2002). Future Warfare and the Decline of Human Decisionmaking. *Parameters* 31 (4), pp. 57–71.

Altmann, J. – Sauer, F. (2017). Autonomous Weapon Systems and Strategic Stability. *Survival* 59 (5), pp. 117–142.

Altmann, J. (2013). Arms control for armed uninhabited vehicles: an ethical issue. *Ethics and Information Technology* 15 (2), pp. 137–152.

Altmann, J. (2009). Preventive Arms Control for Uninhabited Military Vehicles. In: Capurro, R. – Nagenborg, M. eds., *Ethics and Robotics* (Heidelberg: AKA), pp. 69–82.

Altmann, J. (2008). Präventive Rüstungskontrolle. *Die Friedenswarte* 3 (2), pp. 105–126.

Altmann, J. (2004). Military Uses of Nanotechnology: Perspectives and Concerns. *Security Dialogue* 35 (1), pp. 61–79.

Altmann, J. (2001). Non-lethal weapons technologies - the case for independent scientific analysis. *Medicine, Conflict and Survival* 17 (3), pp. 234–247.

Anderson, K. – Reisner, D. – Waxman, M. (2014). Adapting the Law of Armed Conflict to Autonomous Weapon Systems. *International Law Studies* 90 (1), pp. 386–411.

Arkin, R. (2017). A roboticist's perspective on lethal autonomous weapon systems. *UNODA Occasional Papers* 30 (1), pp. 35–48.

Arkin, R. (2010). The Case for Ethical Autonomy in Unmanned Systems. *Journal of Military Ethics* 9 (4), pp. 332–341.

Asaro, P. (2012). On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross* 94 (886), pp. 687–709.

Ashley, R. (1984). The Poverty of Neorealism. *International Organization* 38 (2), pp. 225–286.

Axelrod, R. – Hamilton, W. (1981). The Evolution of Cooperation. *Science* 211 (4489), pp. 1390–1396.

Beier, M. (2017). Short circuit: retracing the political for the age of 'autonomous' weapons. *Critical Military Studies* [forthcoming], pp. 1–18.

Bode, I. – Huelss, H. (2018). Autonomous weapons systems and changing norms in international relations. *Review of International Studies* [forthcoming], pp. 1–21.

Borrie, J. (2017). Security, Unintentional Risk, and System Accidents in the Context of Autonomous Weapon Systems. In: Geiss, R. ed., *Lethal Autonomous Weapons Systems: Technology, Definition, Ethics, Law & Security* (Berlin: Federal Foreign Office), pp. 19–25.

Bull, H. (1976). Arms Control and World Order. *International Security* 1 (1), pp. 3–16.

Cass, K. (2015). Autonomous Weapons and Accountability: Seeking Solutions in the Law of War. *Loyola of Los Angeles Law Review* 48 (1), pp. 1017–1067.

Chengeta, T. (2017). Defining the Emerging Notion of "Meaningful Human Control" in Weapon Systems. *International Law and Politics* 49 (833), pp. 833–890.

Cooper, N. – Mutimer, D. (2011). Arms Control for the 21st Century: Controlling the Means of Violence. *Contemporary Security Policy* 32 (1), pp. 3–19.

Crootof, R. (2016a). The Killer Robots Are Here: Legal and Policy Implications. *Cardozo Law Review* 36 (1), pp. 1837–1915.

Crootof, R. (2016b). A Meaningful Floor for "Meaningful Human Control". *Temple International & Comparative Law Journal* 30 (1), pp. 53–62.

Crootof, R. (2016c). War Torts: Accountability for Autonomous Weapons. *University of Pennsylvania Law Review* 164 (6), pp. 1348–1402.

Davison, N. (2017). A legal perspective: Autonomous weapon systems under international humanitarian law. *UNODA Occasional Papers* 30 (1), pp. 5–18.

Downs, G. – Rocke, D. – Siverson, R. (1985). Arms Races and Cooperation. *World Politics* 38 (1), pp. 118–146.

Ekelhof, M. (2017). Complications of a Common Language: Why it is so Hard to Talk about Autonomous Weapons. *Journal of Conflict & Security Law* 22 (1), pp. 311–331.

Etzioni, A. – Etzioni, O. (2017). Pros and Cons of Autonomous Weapons Systems. *Military Review* (May-June 2017) 72–81.

Fearon, J. (1998). Bargaining, Enforcement, and International Cooperation. *International Organization* 52 (2), pp. 269–305.

Fuhrmann, M. – Horowitz, M. (2017a). Droning On: Explaining the Proliferation of Unmanned Aerial Vehicles. *International Organization* 71 (1), pp. 397 – 418.

Gallarotti, G. (2015). Smart Power: Definitions, Importance, and Effectiveness. Journal of *Strategic Studies* 38 (3), pp. 245–281.

Geist, M. (2016). It's already too late to stop the AI arms race: We must manage it instead. *Bulletin of the Atomic Scientists* 72 (5), pp. 318–321.

Glaser, C. – Kaufmann, C. (1998). What is the Offense-Defense Balance and Can We Measure it? *International Security* 22 (4), pp. 44–82.

Glaser, C. (2003). Structural Realism in a More Complex World. Review of International Studies 29 (3), pp. 403–414.

Glaser, C. (2000). The Causes and Consequences of Arms Races. *Annual Review of Political Science* 3 (1), pp. 251–276.

Glaser, C. (1997). The Security Dilemma Revisited. *World Politics* 50 (1), pp. 171–201.

Glaser, C. (1995). Realists as Optimists: Cooperation as Self-Help. *International Security* 19 (3), pp. 50–90.

Gilpin, R. (1981). *War and Change in World Politics* (New York: Cambridge University Press).

Goldstein, J. – Kahler, M. – Keohane, R. – Slaughter, A. (2000). Introduction: Legalization and World Politics. *International Organization* 54 (3), pp. 385–399.

Goose, S. (2015). The Case for Banning Killer Robots. *Communications of the ACM* 58 (12), pp. 43–45.

Grieco, J. (1988). Realist Theory and the Problem of International Cooperation: Analysis with an Amended Prisoner's Dilemma Model. *The Journal of Politics* 50 (3), pp. 600–624.

Guiora, A. (2017). Accountability and Decision Making in Autonomous Warfare: Who Is Responsible. *Utah Law Review* 2017 (2), pp. 393–422.

Haas, M. – Fischer, S. (2017). The evolution of targeted killing practices: Autonomous weapons, future conflict, and the international order. *Contemporary Security Policy* 38 (2), pp 281–306.

Hadji-Janev, M. – Hristovski, K. (2017). Beyond the Fog: Autonomous Weapon Systems in the Context of the International Law of Armed Conflicts. *Jurimetrics* 57 (1), pp. 325–340.

Hamlett, P. (1990). Technology and the Arms Race. *Science, Technology, & Human Values* 15 (4), pp. 461–473.

Hansel, M. – Mutschler, M. – Dickow, M. (2018). Taming cyber warfare: lessons from preventive arms control. *Journal of Cyber Policy* [forthcoming], pp. 1–17.

Herz, J. (1950). Idealist Internationalism and the Security Dilemma. *World Politics* 2 (2), pp. 157–180.

Hoffman, R. – Cullen, T. – Hawley, J. (2016). The myths and costs of autonomous weapon systems. *Bulletin of the Atomic Scientists* 72 (4), pp. 247–255.

Hollyer, J. – Rosendorff, P. – Vreeland, J. (2011). Democracy and Transparency. *The Journal of Politics* 73 (4), pp. 1191–1205.

Horowitz, M. (2016). Public opinion and the politics of the killer robots debate. *Research and Politics* 3 (1), pp. 1–8.

Huntington, S. (1958). Arms Races: Prerequisites and Results. *Public Policy* 8 (1), pp. 41-86.

Jenks, C. (2016). False Rubicons, Moral Panic, & Conceptual Cul-De-Sac: Critiquing & Reframing the Call to Ban Lethal Autonomous Weapons. *Pepperdine Law Review* 44 (1), pp. 1–70.

Jervis, R. (1999). Realism, Neoliberalism, and Cooperation: Understanding the Debate. *International Security* 24 (1), pp. 42–63.

Jervis, R. (1993). Arms Control, Stability, and Causes of War. *Political Science Quarterly* 108 (2), pp. 239–253.

Jervis, R. (1988). Realism, Game Theory, and Cooperation. *World Politics* 40 (3), pp. 317–349.

Jervis, R. (1978). Cooperation Under the Security Dilemma. *World Politics* 30 (2), pp. 167–214.

Karppi, T. – Böhlen, M. – Granata, Y. (2018). Killer Robots as cultural techniques. *International Journal of Cultural Studies* 21 (2), pp. 107–123.

Koremenos, B. (2001). Loosening the Ties That Bind: A Learning Model of Agreement Flexibility. *International Organization* 55 (1), pp. 289-325.

Kreps, S. (2018). The Institutional Design of Arms Control Agreements. *Foreign Policy Analysis* 14 (1), pp 127–147.

Kydd, A. (2000). Arms Races and Arms Control: Modelling the Hawk Perspective. *American Journal of Political Science* 44 (2), pp. 228–244.

Lachow, I. (2017). The upside and downside of swarming drones. *Bulletin of the Atomic Scientists* 73 (2), pp. 96–101.

Lakoff, S. – Bruvold, E. (1990). Controlling the Qualitative Arms Race: The Primacy of Politics. *Science, Technology, & Human Values* 15 (4), pp. 382–411.

Lebow, R. (1996). Thomas Schelling and Strategic Bargaining. *International Journal* 51 (3), pp. 555–576.

Legro, J. – Moravcsik, A. (1999). Is Anybody Still a Realis? *International Security* 24 (2), pp. 5–55.

Levy, J. (1984). The Offensive/Defensive Balance of Military Technology: A Theoretical and Historical Analysis. *International Studies Quarterly* 28 (2), pp. 219–238.

Leys, N. (2018). Autonomous Weapon Systems and International Crises. *Strategic Studies Quarterly* 12 (1) pp. 48–73.

Lin, P. (2010). Ethical Blowback from Emerging Technologies. *Journal of Military Ethics* 9 (4), pp. 313–331.

MacKenzie, D. (1989). Review: Technology and the Arms Race. *International Security* 14 (1), pp. 161–175.

Marra, W. – McNeil, S. (2013). Understanding 'The Loop': Regulating the Next Generation of War Machines. *Harvard Journal of Law and Public Policy* 36 (3), pp. 1–62.

Marchant, G. – Allenby, B. – Arkin, R. – Barrett, E. – Borenstein, J. – Gaudet, L. – Kittrie, O. – Lin, P. – Lucas, G. – O'Meara, R. – Silberman, J. (2011). International Governance of Autonomous Military Robots. *The Columbia Science and Technology Law Review* 12 (1), pp. 272–315.

Mearsheimer, J. (1995). The False Promise of Internatinal Institutions. *International Security* 19 (3), pp. 5–49.

McCubbins, M. (1983). Policy Components of Arms Competition. *American Journal of Political Science* 27 (3), pp. 385–406.

Midgaard, K. (1970). Arms Races, Arms Control, and Disarmament. *Cooperation and Conflict* 5 (1), pp. 20–51.

Morrow, J. (1994a). Modeling the Forms of International Cooperation: Distribution Versus Information. *International Organization* 48 (3), pp. 387–423.

Morrow, J. (1989). A Twist of Truth: A Reexamination of the Effects of Arms Races on the Occurrence of War. *The Journal of Conflict Resolution* 33 (3), pp. 500–529.

Neslage, K (2016). Does "Meaningful Human Control" Have Potential for the Regulation of Autonomous Weapon Systems? *University of Miami National Security & Armed Conflict Law Review* 6 (1), pp. 151–177.

Nye, J. (2011). Power and foreign policy. *Journal of Political Power* 4 (1), pp. 9–24.

O'Connell, M. (2015). 21st Century Arms Control Challenges: Drones, Cyber Weapons, Killer Robots and WMDs. *Washington University Global Studies Law Review* 13 (515), pp. 515–533.

Oye, K. (1985). Explaining Cooperation under Anarchy: Hypotheses and Strategies. *World Politics* 38 (1), pp. 1–24.

Poast, P. (2012). Do Issue Linkage Work? Evidence from European Alliance Negotiations, 1860 to 1945. *International Organization* 66 (2), pp. 277–310.

Putnam, R. (1988). Diplomacy and Domestic Politics: The Logic of Two-Level Games. *International Organization* 42 (3), pp 427–460.

Ramlo, S. (2011). Using word clouds to present Q methodology data findings. *Human Subjectivity* 9 (2), pp. 99–111.

Rathbun, B. (2008). A Rose by Any Other Name: Neoclassical Realism as the Logical and Necessary Extension of Structural Realism. *Security Studies* 17 (1), pp. 294–321.

Roff, H. (2014). The Strategic Robot Problem: Lethal Autonomous Weapons in War. *Journal of Military Ethics* 13 (3), pp. 211–227.

Rothman, S. (2011). Revising the soft power concept: what are the means and mechanisms of soft power? *Journal of Political Power* 4 (1), pp. 49–64.

Sehrawat, V. (2017). Autonomous weapon system: Law of armed conflict (LOAC) and other legal challenges. *Computer Law & Security Review* 33 (1), pp. 38–56.

Scharre, P. (2017). A security perspective: Security concerns and possible arms control approaches. *UNODA Occasional Papers* 30 (1), pp. 19–34.

Scharre, P. (2015). The Opportunity and Challenge of Autonomous Systems. In: Williams, A. - Scharre, P. eds., *Autonomous Systems: Issues for Defence Policymakers* (The Hague: NATO Communications and Information Agency), pp. 3–26.

Schelling, T. (1956). An Essay on Bargaining. *The American Economic Review* 46 (3), pp. 281–306.

Schmitt, M. – Thurnher, J. (2013). "Out of the Loop": Autonomous Weapon Systems and the Law of Armed Conflict. *Harvard National Security Journal* 4 (1), pp. 231–281.

Schofield, J. (2000). Arms Control Failure and the Balance of Power. *Canadian Journal of Political Science* 33 (4), pp. 747–777.

Schweller, R. – Priess, D. (1997). A Tale of Two Realisms: Expanding the Institutions Debate. *Mershon International Studies Review* 41 (1), pp. 1–32.

Sharkey, N. (2017). Why robots should not be delegated with the decision to kill. *Connection Science* 29 (2), pp. 177–186.

Sharkey, N. (2012a). The evitability of autonomous robot warfare. *International Review of the Red Cross* 94 (886), pp. 787–799.

Sharkey, N. (2010). Saying 'No!' to Lethal Autonomous Targeting. *Journal of Military Ethics* 9 (4), pp. 369–383.

Slapin, J. – Proksch, S. (2008). A Scaling Model for Estimating Time-Series Party Positions from Texts. *American Journal of Political Science* 52 (3), pp. 705–722.

Sparrow, R. (2007). Killer Robots. *Journal of Applied Philosophy* 24 (1), pp. 62–77.

Steinberg, G. (1985). The Role of Process in Arms Control Negotiations. *Journal of Peace Research* 2 (3), pp. 261–272.

Sterling-Folker, J. (1997). Realist Environment, Liberal Process, and Domestic-Level Variables. *International Studies Quarterly* 41 (1), pp. 1–25.

Taliaferro, J. (2000). Security Seeking under Anarchy: Defensive Realism Revisited. *International Security* 25 (3), pp. 128–161.

Tang, S. (2010). Offence-defence Theory: Towards a Definitive Understanding. *The Chinese Journal of International Politics* 3 (1), pp. 213–260.

Tang, S. (2009). The Security Dilemma: A Conceptual Analysis. *Security Studies* 18 (3), pp. 587–623.

Theodorou, A. – Wortham, R. – Bryson, J. (2017). Designing and implementing transparency for real time inspection of autonomous robots. *Connection Science* 29 (3), pp. 230–241.

Van Evera, S. (1998). Offense, Defense, and the Causes of War. *International Security* 22 (4), pp. 5–43.

Wallach, W. (2017b). Toward a Ban on Lethal Autonomous Weapons: Surmounting the Obstacles. *Communications of the ACM* 60 (5), pp. 28–34.

Walt, S. (1998). International Relations: One World, Many Theories. *Foreign Policy* 110 (1), pp. 29–46.

Walt, S. (1985). Alliance Formation and the Balance of World Power. *International Security* 9 (4), pp. 3–43.

Wendt, A. (1992). Anarchy is what States Make of it: The Social Construction of Power Politics. *International Organization* 46 (2), pp. 391–425.

Wendt, A. (1987). The Agent-Structure Problem in International Relations Theory. *International Organization* 41 (3), pp. 335–370.

Wilson, E. (2008). Hard Power, Soft Power, Smart Power. *The Annals of the American Academy of Political and Social Science* 616 (1), pp. 110–124.

Zakaria, F. (1992). Reealism and Domestic Politics: A Review Essay. *International Security* 17 (1), pp. 177–198.

8.3. Online Sources

Anderson, K. – Waxman, M. (2013). Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can. *Hoover Institution* (available at: https://goo.gl/A66zEg, 3. 5. 2018).

Arkin, R. – Ulam, P. – Duncan, B. (2009). An Ethical Governor for Constraining Lethal Action in an Autonomous System. *Georgia Institute of Technology, Mobile Robot Lab, College of Computing, Technical Report* (available at: https://goo.gl/tQWiBh, 16. 1. 2018).

Article 36 (2014). Key areas for debate on autonomous weapons systems. *Article 36, Briefing Paper* (available at: https://goo.gl/Mpsp3E, 18. 1. 2018).

CSKR (2018a). Chronology. *Campaign to Stop Killer Robots* (available at: https://www.stopkillerrobots.org/chronology/, 20. 7. 2018).

CSKR (2018b). *Country Views on Killer Robots* (available at: https://goo.gl/rpYSnW, 21. 7. 2018).

Docherty, B. (2014). Taking on “Killer Robots”. *Just Security* (available at: https://goo.gl/nmp2oi, 19. 7. 2018).

DoD (2013). *Unmanned Systems Integrated Roadmap FY2013-2038* (available at: https://goo.gl/yvGmou, 27. 3. 2018).

DoD (2012). *Directive 3000.09* (available at: https://cryptome.org/dodi/dodd-3000-09.pdf, 23. 3. 2018).

DoD (2007). *Unmanned Systems Roadmap 2007-2032* (available at: https://goo.gl/cWAQPe, 1. 3. 2018).

Dunkelberg, A. (2016). *Laws for L.A.W.S.: Legal Challenges for the Use of Lethal Autonomous Weapons Systems in Times of Armed Conflict* (available at: https://goo.gl/77nf5B, 28. 1. 2018).

Ewers, E. – Fish, L. – Horowitz, M. – Sander, A. – Scharre, P. (2017). Drone Proliferation: Policy Choices for the Trump Administration. *CNAS* (available at: https://goo.gl/KR1Ft9, 4. 3. 2018).

FLI (2017). An Open Letter to the United Nations Convention on Certain Conventional Weapons. *Future of Life Institute* (available at: https://goo.gl/yhU4h1, 28. 1. 2018).

FLI (2016). Benefits & Risks of Artificial Intelligence. *Future of Life Institute* (available at: https://goo.gl/zPCVGX, 27. 3. 2018).

FLI (2015). Autonomous Weapons: An Open Letter From AI & Robotics Researchers. *Future of Life Institute* (available at: https://futureoflife.org/open-letter-autonomous-weapons/, 28. 1. 2018).

Geiss, R. (2015). The International-Law Dimension of Autonomous Weapons Systems. *Friendrich Ebert Stiftung* (available at: http://library.fes.de/pdf-files/id/ipa/11673.pdf, 29. 1. 2018).

Greig, R. – Enterline, A. (2010). Correlates of War Project: National Material Capabilities Data Documentation. *COW* (available at: https://goo.gl/mhhsTA. 19. 7. 2018).

Gubrud, M. – Altmann, J. (2013). Compliance Measures for an Autonomous Weapons Convention. *ICRAC Working Paper* (available at: https://goo.gl/ru3uCJ, 6. 5. 2018).

Heyns, C. (2013). Report of the Special Rapporteur on extrajudicial, summary or arbitrary executions, Christof Heyns. *Human Rights Council, 23rd Session* (available at: https://goo.gl/OJ8TXI, 19. 1. 2018).

HRW (2012). Losing Humanity: Case against Killer Robots. *Human Rights Watch* (available at: https://goo.gl/kl2qNI, 17. 1. 2018).

Horowitz, M. – Scharre, P. (2015). Meaningful Human Control in Weapon Systems: A Primer. *CNAS* (available at: https://goo.gl/yy8rGK, 3. 5. 2018).

ICRC (2016). Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons. *Expert Meeting Report* (available at: https://goo.gl/UyioAB, 23. 3. 2018).

ICRC (2014). Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects. *Expert Meeting Report* (available at: https://goo.gl/iM2zkL, 3. 3. 2018).

ILW (2017). Integrating Army Robotics and Autonomous Systems to Fight and Win. *Institute of Land Warfare* (available at: https://goo.gl/MohEgz, 23. 2. 2018).

Investopedia (2018). *Multiple Linear Regression - MLR* (available at: https://www.investopedia.com/terms/m/mlr.asp, 29. 7. 2018).

Ipsos (2017). *Three in Ten Americans Support Using Autonomous Weapons* (available at: https://goo.gl/xmYkJW, 1. 3. 2018).

Jenks, C. – Ford, C. (2016). The International Discussion Continues: 2016 CCW Experts Meeting on Lethal Autonomous Weapons. *Just Security* (available at: https://goo.gl/yE9nPb, 19. 7. 2018).

Kania, E. (2017). Battlefield Singularity: Artificial Intelligence, Military Revolution, and China's Future Military Power. *Center for New American Security* (available at: https://goo.gl/RS37PJ, 5. 7. 2018).

Lane, D. (2018). Log Transformations. *OnlineStatBook* (available at: https://goo.gl/4wDpXQ, 29. 7. 2018).

Lin, P. – Bekey, G. – Abney, K. (2008). Autonomous Military Robotics: Risk, Ethics, and Design. *US Department of Navy* (available at: http://ethics.calpoly.edu/onr_report.pdf, 26. 2. 2018).

Neuneck, G. & Mölling, C. (2001). Methoden, Kriterien und Konzepte für Präventive Rüstungskontrolle. *Wissenschaft und Frieden* (available at: https://goo.gl/ahqBRm, 14. 7. 2018).

NDU (2017). Spring 2017 Industry Study, Final Report: Robotics and Autonomous Systems. *The Dwight D. Eisenhower School for National Security and Resource Strategy, National Defense University* (available at: https://goo.gl/kK6GVP, 1. 3. 2018).

ORi (2015). *The Ethics and Governance of Lethal Autonomous Weapons Systems: An International Public Opinion Poll.* Open Roboethics initiative (available at: https://goo.gl/hPYfGX, 1. 3. 2018).

Reaching Critical Will (2018). *UN General Assembly First Committee* (available at: http://reachingcriticalwill.org/disarmament-fora/unga, 21. 7. 2018).

Roff, H. (2016). Weapons autonomy is rocketing. *Foreign Policy* (available at: https://goo.gl/F5biHa, 2. 3. 2018).

Roff, H. – Danks, D. (2017). *"Trust but Verify": The difficulty of trusting autonomous weapons systems* (available at: https://goo.gl/hPQmh4, 4. 3. 2018).

Roff, H. (2017). What Do People Around the World Think About Killer Robots? International law about autonomous weapons systems needs to consider public opinion. *Slate* (available at: https://goo.gl/eoBuHq, 1. 3. 2018).

Roff, H. – Moyes, R. (2016). Meaningful Human Control, Artificial Intelligence and Autonomous Weapons. *Article 36* (available at: https://goo.gl/3w38D6, 3. 5. 2018).

Rosert, E. (2017). How to Regulate Autonomous Weapons: Steps to Codify Meaningful Human Control as a Principle of International Humanitarian Law.

Peace Research Institute Frankfurt (available at: https://goo.gl/4dyGid, 18. 1. 2018).

Scharre, P. (2016). Autonomous Weapons and Operational Risk. *CNAS* (available at: https://goo.gl/NtgJ9Q, 5. 3. 2018).

Scharre, P. – Horrowitz, M. (2015). An Introduction to Autonomy in Weapon Systems. *CNAS* (available at: https://goo.gl/fPi1sd, 20. 3. 2018).

Shachtman, N. (2007). Robot Cannon Kills 9, Wounds 14. *Wired* (available at: https://goo.gl/SsLafd, 4. 3. 2018).

Silge, J. – Robinson, D. (2018). Relationships between words: n-grams and correlations. *Text Mining with R: A Tidy Approach* (available at: https://www.tidytextmining.com/ngrams.html, 27. 7. 2018).

Tucker, P. (2017). Russia to the United Nations: Don't Try to Stop Us From Building Killer Robots. *Defense One* (available at: https://goo.gl/A73twq, 28. 7. 2018).

UNIDIR (2017a). The Weaponization of Increasingly Autonomous Technologies: Autonomous Weapon Systems and Cyber Operations. *The United Nations Institute for Disarmament Research* (available at: https://goo.gl/nS3zKA, 5. 3. 2018).

UNIDIR (2017b). The Weaponization of Increasingly Autonomous Technologies: Concerns, Characteristics and Definitional Approaches. *The United Nations Institute for Disarmament Research* (available at: https://goo.gl/Z7WtA4, 6. 3. 2018).

UNIDIR (2016). Safety, Unintentional Risk and Accidents in the Weaponization of Increasingly Autonomous Technologies. *The United Nations Institute for Disarmament Research* (available at: https://goo.gl/Uwkq8s, 3. 3. 2018).

UNIDIR (2015). The Weaponization of Increasingly Autonomous Technologies: Considering Ethics and Social Values. *The United Nations Institute for Disarmament Research* (available at: https://goo.gl/izKDSk, 3. 3. 2018).

UNIDIR (2014). Framing Discussions on the Weaponization of Increasingly Autonomous Technologies. *The United Nations Institute for Disarmament Research* (available at: https://goo.gl/sZPVJY, 3. 3. 2018).

UNOG (2018). *Discussions on emerging technologies in the area of LAWS* (available at: https://goo.gl/C3EXmL, 21. 7. 2018).

US Army (2017). *The US Army Robotic and Autonomous Systems Strategy* (available at: https://goo.gl/cLBKTk, 26. 2. 2018).

Wezeman, P. – Fleurant, A. – Kuimova, A. – Tian, N. – Wezeman, T. (2018). Trends in International Arms Transfers, 2017. *SIPRI* (available at: https://goo.gl/SU4U8d, 20. 3. 2018).

8.4. Data Sources

Brazilian Delegation (2013). Statement by Brazil at Human Rights Council 23rd Session with Special Rapporteur on extrajudicial, summary or arbitrary executions. *CSKR* (available at: https://goo.gl/d3wyVw, 27. 7. 2018).

Brazilian Delegation (2014). Statement by H.E. Ambassador Pedro Motta Pinto Coelho Permanent Representative of Brazil to the Conference on Disarmament at CCW Informal Expert Meeting on Lethal Autonomous Systems, Geneva 13 May 2014. *Reaching Critical Will* (available at: https://goo.gl/DMsWtR, 27. 7. 2018).

Center for Systemic Peace (2018a). *Major Episodes of Political Violence, 1946-2017* (available at: http://www.systemicpeace.org/inscrdata.html, 29. 7. 2018).

Center for Systemic Peace (2018b). *Polity IV Annual Time-Series, 1800-2017* (available at: http://www.systemicpeace.org/inscrdata.html, 29. 7. 2018).

Chinese Delegation (2013). Intervention at the Meeting of the High Contracting Parties to the CCW. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Chinese Delegation (2016a). Statement by H.E. Ambassador Fu Cong At Preparatory Committee for the 5th Review Conference to the Convention on Certain Conventional Weapons. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Chinese Delegation (2016b). Statement by Ambassador WANG Qun, Director-General of the Arm Control Department of the Ministry of Foreign Affairs of China, at the General Debate of the First Committee of the 71st Session of the UNGA. *Reaching Critical Will* (available at: https://goo.gl/DMsWtR, 27. 7. 2018).

French Delegation (2014a). Convention sur Certaines Armes Classiques (CCAC), Réunion des Etats parties (Geneve, 13-14 novembre 2014), Systemes d'armes létaux autonomes, Intervention nationale. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2014b). Convention sur Certaines Armes Classiques (CCAC), Réunion informelle d'experts sur les systemes d'armes létaux autonomes (SALA) (Geneve, 13-16 mai 2014), Intervention générale. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2014c). Convention sur Certaines Armes Classiques (CCAC), Réunion informelle d'experts sur les systemes d'armes létaux autonomes (SALA) (Geneve, 13-16 mai 2014), Intervention de M. Jean-Hughes SIMON-MICHEL, Ambassadeur Représentant Permanent de la France - Aspects techniques des SALA. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2015a). Convention sur Certaines Armes Classiques (CCAC), Réunion des hautes parties contractantes a la CCA, Geneve, 12-13 novembre 2015, Intervention générale. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2015b). Convention sur Certaines Armes Classiques (CCAC), Réunion des Etats parties (Geneve, 12-13 novembre 2015), Intervention Systemes d'armes létaux autonomes (SALA). *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2015c). Convention sur Certaines Armes Classiques (CCAC), Réunion informelle d'experts sur les systemes d'armes létaux autonomes (SALA) (Geneve, 13-17 avril 2015), Intervention générale. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2015d). Convention sur Certaines Armes Classiques (CCAC), Réunion informelle d'experts sur les systemes d'armes létaux autonomes (SALA)

(Geneve, 13-17 avril 2015), Aspects éthiques des SALA. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2016a). Convention on Certain Conventional Weapons (CCW), Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS) (Geneva, 11-15 April 2016), non-paper on mapping of technological development. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2016b). Convention on Certain Conventional Weapons (CCW), Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Geneva, 11-15 April 2016, non-paper on legal framework for any potential development and operational use of a future lethal autonomous weapons system (LAWS). *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2016c). Convention on Certain Conventional Weapons (CCW), Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS), Geneva, 11-15 April 2016, non-paper on characterization of a LAWS. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2016d). Convention sur Certaines Armes Classiques (CCAC), Réunion d'experts sur les systemes d'armes létaux autonomes (SALA) (Geneve, 11-15 avril 2016), « Vers une définition opérationnelle des SALA » . *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

French Delegation (2017). For Consideration by the Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS) - paper submitted by France and Germany. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Fuhrmann, M. – Horowitz, M. (2017b). *Supplementary materials* (available at: https://goo.gl/7kUNk6, 29. 7. 2018).

German Delegation (2014a). CCW Expert Meeting on Lethal Autonomous Weapon Systems, Geneva, 13-16 May 2014, General Statement by Germany. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2014b). CCW Expert Meeting on Lethal Autonomous Weapon Systems, Geneva, 13-16 May 2014, Statement on Ethics and Sociology of LAWS by Germany. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2014c). CCW Expert Meeting on Lethal Autonomous Weapon Systems (LAWS), Geneva 13-16 May, Statement on Legal Aspects by Ambassador Michael Biontino, May 14, 2014. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2014d). CCW Expert Meeting on Lethal Autonomous Weapon Systems, Geneva, 13-6 May 2014, Statement Wrap-up Session by Michael Biontino. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2015a). CCW Expert Meeting on Lethal Autonomous Weapon Systems, Geneva, 13-17 April 2014, General Statement by Germany. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2014b). CCW Expert Meeting on Lethal Autonomous Weapon Systems (LAWS), Geneva 13-16 May, Statement on Legal Aspects by Ambassador Michael Biontino, May 14, 2014. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2017a). For Consideration by the Group of Governmental Experts on Lethal Autonomous Weapons Systems (LAWS) - paper submitted by France and Germany. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

German Delegation (2017b). General Statement by H.E. Ambassador Michael Biontino, Permanent Representative of Germany to the Conference on Disarmament and Ambassador for Global Disarmament Affairs, First Meeting of the Group of Governmental Experts on Lethal Autonomous Weapons Systems, Geneva, 13 to 17 November 2017. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

In.europa (2017). *State Power Index* (available at: http://index.ineuropa.pl/en/state-power-index/, 21. 7. 2018).

Indian Delegation (2014a). Statement by Ambassador D.B. Venkatesh Varma, Permanent Representative of India to the CD, at the 2014 Meeting of HCPs to the CCW, Geneva, November 13, 2014. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Indian Delegation (2014b). Intervention by Indian Delegation at the Meeting of the High Contracting Parties. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Indian Delegation (2015a). Statement by Ambassador D.B. Venkatesh Varma, Permanent Representative of India to the CD at the 2015 Meeting of High Contracting Parties to the CCW, November 12, 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Indian Delegation (2015b). Statement by Ambassador D.B. Venkatesh Varma, Permanent Mission of India to the CD at the CCW Informal Meeting of Experts on Lethal Autonomous Weapons Geneva 17 April 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Indian Delegation (2016a). Statement by Ambassador D.B. Venkatesh Varma, Permanent Representative of India to the CD at the Preparatory Committee for the Fifth Review Conference of High Contracting Parties to the CCW, Geneva, August 31, 2016. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Indian Delegation (2016b). Statement by Ambassador D.B. Venkatesh Varma, Permanent Mission of India to the CD at the CCW Informal Meeting of Experts on Lethal Autonomous Weapons, 11 April 2016. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Indian Delegation (2017). Statement by Ambassador Amandeep Singh Gill, Permanent Representative of India to the Conference on Disarmament at the Thematic Debate on Other Disarmament Measures and International Security, First Committee, 72nd Session of the United Nations General Assembly, New York, 23 October 2017. *Reaching Critical Will* (available at: https://goo.gl/DMsWtR, 27. 7. 2018).

Israeli Delegation (2014). Statement by Israel by Mr. Eitan Levon, Representative to the Conference on Disarmament, Meeting of the High Contracting Parties to the Convention on Certain Conventional Weapons (CCW), 13 November 2014. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Israeli Delegation (2015a). Statement by Mrs. Maya Yaron, Counsellor Deputy Permanent Representative to the Conference on Disarmament at the Meeting of the High Contracting Parties to the Convention on Certain Conventional Weapons (CCW) - General Exchange of Views, Geneva, 12 November 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Israeli Delegation (2015b). Statement by Israel by Mr. Eitan Levon, Representative to the Conference on Disarmament at the Meeting of Experts on Lethal Autonomous Weapon Systems (LAWS), 13 April 2015 - General Debate. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Israeli Delegation (2016). Statement on Lethal Autonomous Weapons Systems (LAWS) by Mrs. Maya Yaron, Counsellor Deputy Representative to the Conference on Disarmament, Group of Experts meeting on Lethal Autonomous Weapons Systems (LAWS), April 11-15 2016, Geneva - General Debate. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Italian Delegation (2016). CCW Meeting of Experts on Lethal Autonomous Weapons Systems, Geneva (11-15 April 2016), Statement by the Delegation of Italy to the Conference on Disarmament - Towards a working definition of LAWS. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Japanese Delegation (2014). Statement by H.E. Ambassador Toshio SANO, Permanent Representative of Japan to the Conference on Disarmament, Experts Meeting on Lethal Autonomous Weapons Systems to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects, 13 May 2014, Geneva. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Japanese Delegation (2015). Japan’s views on issues relating to LAWS. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Korean Delegation (2014). Statement by Ambassador AHN Youngjip at the Meeting of Experts on Lethal Autonomous Weapons Systems, Convention on Certain Conventional Weapons, Geneva 13 May 2014. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Korean Delegation (2015a). Statement by Ambassador AHN Youngjup at the Meeting of Experts on Lethal Autonomous Weapons Systems, Convention on Certain Conventional Weapons, Geneva 13 April 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Korean Delegation (2015b). Statement by H.E. Ambassador KIM Young-moo of the Republic of Korea at the 2015 Meeting of the High Contracting Parties to the CCW (Lethal Autonomous Weapons Systems). *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Mijares, V. (2016). *CINC Dataset Updated: 2008-2013* (available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2783521, 28. 4. 2018).

Russian Delegation (2014a). Statement by Mr. Vladimir Yermakov, Head of Delegation of the Russian Federation, Deputy Director of the Department for Non-Proliferation and Arms Control, Ministry of Foreign Affairs of the Russian Federation at the Meeting of High Contracting Parties to the Convention on Certain Conventional Weapons in Geneva, 13 November 2014. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Russian Delegation (2014b). Intervention by Russian Delegation at the Meeting of the High Contracting Parties to the CCW. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

Russian Delegation (2017). Examination of various dimensions of emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and purposes of the Convention. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

SIPRI (2018). *SIPRI Military Expenditure Database* (available at: https://www.sipri.org/databases/milex, 29. 7. 2018).

UK Delegation (2015a). UK Statement to the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, 13-17th April 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

UK Delegation (2015b). UK Statement on Lethal Autonomous Weapons Systems to the CCW Meeting of the High Contracting Parties, 12-13th November 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

UK Delegation (2016a). UK Statement to the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, 11-15 April 2016 - General Debate. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

UK Delegation (2016b). UK Statement to the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, 11-15 April 2016 - Possible Challenges to IHL due to Increasing Degrees of Autonomy. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

UK Delegation (2016c). UK Statement to the Informal Meeting of Experts on Lethal Autonomous Weapons Systems, 11-15 April 2016 - Working towards a definition of LAWS. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

UNCTAD (2018). *UN list of Least Developed Countries* (available at: https://goo.gl/bUUSRc, 29. 7. 2018).

US Delegation (2013). Clustered Interactive Dialogue on Extrajudicial, Summary, and Arbitrary Killing and Internally Displaced Persons, Statement by the Delegation of the United State of America at Human Rights Council - 23rd Session in Geneva, May 29, 2013. *CSKR* (available at: https://goo.gl/d3wyVw, 27. 7. 2018).

US Delegation (2014a). Meeting of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

US Delegation (2014b). U.S. Delegation Opening Statement at the 2014 Meeting of Experts on Lethal Autonomous Weapons Systems. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

US Delegation (2014c). Intervention by US Delegation at the Meeting of Experts on Lethal Autonomous Weapons Systems. *UNOG* (available at: https://goo.gl/UEVZ wm, 27. 7. 2018).

US Delegation (2015). The Convention on Certain Conventional Weapons (CCW) Informal Meeting of Experts on Lethal Autonomous Weapons Systems, U.S. Delegation Opening Statement as Delivered by Michael W. Meier, Geneva, April 13, 2015. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

US Delegation (2017a). Autonomy in Weapon Systems - Paper Submitted by the United States of America. *UNOG* (available at: https://goo.gl/UEVZwm, 27. 7. 2018).

US Delegation (2017b). Characteristics of Lethal Autonomous Weapons Systems - Paper Submitted by the United States of America. *UNOG* (available at: https://goo. gl/UEVZwm, 27. 7. 2018).

WB (2018a). *GDP per capita (current US$)* (available at: https://goo.gl/Py4Ehz, 29. 7. 2018).

WB (2018b). *Armed forces personnel, total* (available at: https://goo.gl/S4Tr5N, 29. 7. 2018).

WB (2018c). *Population, total* (available at: https://goo.gl/n9mxAw, 29. 7. 2018).

WB (2018d). *Research and development expenditure (% of GDP)* (available at: https://goo.gl/jrkhBY, 29. 7. 2018).

WB (2018e). *GDP (current US$)* (available at: https://goo.gl/RZFtqd, 29. 7. 2018).

WB (2018f). *High-technology exports (current US$)* (available at: https://data. worldbank.org/indicator/TX.VAL.TECH.CD?end=2016&start=2000, 22. 7. 2018).

9. Appendices

Appendix 1: Criteria of Preventive Arms Control

Criteria of Preventive Arms Control:

I. *Adherence to and further development of effective arms control, disarmament and international law*
   1. *Prevent dangers to existing or intended arms-control and disarmament treaties*
   2. *Observe existing norms of humanitarian law*
   3. *No utility for weapons of mass destruction*

II. *Maintain and improve stability*
   1. *Prevent destabilization of the military situation*
   2. *Prevent technological arms race*
   3. *Prevent horizontal or vertical proliferation/diffusion of military-relevant technologies, substances or knowledge*

III. *Protect humans, environment and society*
   1. *Prevent dangers to humans*
   2. *Prevent dangers to the environment and sustainable development*
   3. *Prevent dangers to the development of societal and political systems*
   4. *Prevent dangers to the societal infrastructure*

Source: Altmann (2009: 73–74)

Appendix 2: Formulas for Measures of Hard and Soft Power

CINC formula:

$$CINC = \frac{TPR + UPR + ISPR + ECR + MER + MPR}{6}$$

Where:

TPR – Total Population Ratio

UPR – Urban Population Ratio

$ISPR$ – Iron and Steel Production Ratio

ECR – Primary Energy Consumption Ratio

MER – Military Expenditure Ratio

MPR – Military Personnel Ratio

Source: Mijares (2016)

SPI formula:

$$IMP_{ik} = IKa_{ik} \times 40\% + IM_{ik} \times 20\% + IZ_{ik} \times 10\% + IP_{ik} \times 10\% + IKu_{ik} \times 10\% + IZN_{ik} \times 5\% + ID_{ik} \times 5\%$$

Where:

IKa_{ik} – Capital Index

IM_{ik} –Militarization Index

IZ_{ik} – Land Index

IP_{ik} – Population Index

IK_{ik} – Culture Index

IZN_{ik} – Natural Resources Index

ID_{ik} – Diplomacy Index

Source: In.europa (2017)

CCINC formula

$$CCINC = \frac{TPR + UPR + ISPR + ECR + MER + MPR + 2 \times (IBAWS)}{8}$$

Where:

TPR – Total Population Ratio

UPR – Urban Population Ratio

$ISPR$ – Iron and Steel Production Ratio

ECR – Primary Energy Consumption Ratio

MER – Military Expenditure Ratio

MPR – Military Personnel Ratio

$IBAWS$ – Industrial Base for Autonomous Weapon Systems

Source: Author's elaboration based on Mijares (2016)

CSPI formula:

$$IMP_{ik} = IKa_{ik} \times 32\% + IM_{ik} \times 16\% + IZ_{ik} \times 8\% + IP_{ik} \times 8\% + IKu_{ik} \times 8\% + IZN_{ik} \times 4\% + ID_{ik} \times 4\% + IBAWS \times 20\%$$

Where:

IKa_{ik} – Capital Index

IM_{ik} –Militarization Index

IZ_{ik} – Land Index

IP_{ik} – Population Index

IK_{ik} – Culture Index

IZN_{ik} – Natural Resources Index

ID_{ik} – Diplomacy Index

$IBAWS$ – Industrial Base for Autonomous Weapon Systems

Source: Author's elaboration based on In.europa (2017)

IBAWS formula:

$$IBAWS = RDR \times 50\% + AER \times 25\% + HTER \times 25\%$$

Where:

RDR – R&D Spending Ratio

AER – Arms Exports Ratio

$HTER$ – High-Tech Exports Ratio

Source: Author (2018)

Appendix 3: Order of Preferences

Group 1

BRA: Partial Prohibition > Total Prohibition > Hybrid Reg. > Minor Reg. > No Reg.

IND: Hybrid Reg. > Partial Prohibition > Minor Reg. > Total Prohibition > No Reg.

CHN: Hybrid Reg. > Partial Prohibition >Minor Reg. >Total Prohibition > No Reg.

Group 2

DEU: Hybrid Reg. > Partial Prohibition > Minor Reg. > Total Prohibition > No Reg.

GBR: Hybrid Reg. > Minor Reg. > Partial Prohibition > No Reg. > Total Prohibition

JPN: Hybrid Reg. > Minor Reg. > Partial Prohibition > No Reg. > Total Prohibition

FRA: Hybrid Reg. > Minor Reg. > Partial Prohibition > No Reg. > Total Prohibition

ITA: Hybrid Reg. > Minor Reg. > Partial Prohibition > No Reg. > Total Prohibition

KOR: Minor Reg. > Hybrid Reg. > No Reg. >Partial Prohibition > Total Prohibition

Group 3

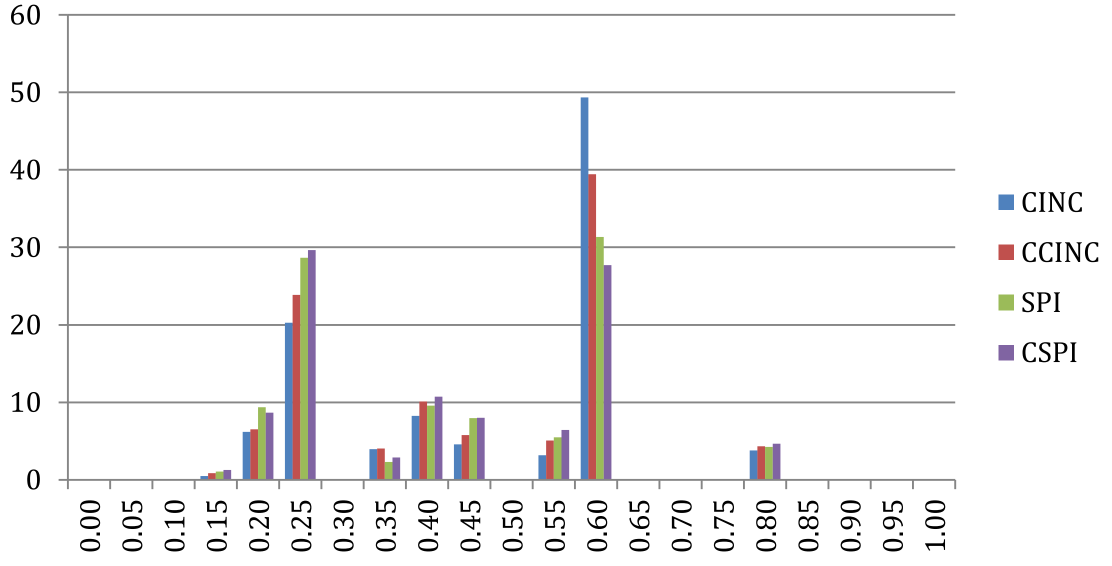
USA: Minor Reg. > No Reg. > Hybrid Reg. > Partial Prohibition > Total Prohibition

RUS: Minor Reg. > No Reg. > Hybrid Reg. > Partial Prohibition > Total Prohibition

ISR: Minor Reg. > No Reg. > Hybrid Reg. > Partial Prohibition > Total Prohibition

Appendix 4: Distribution of Power across the Policy Continuum

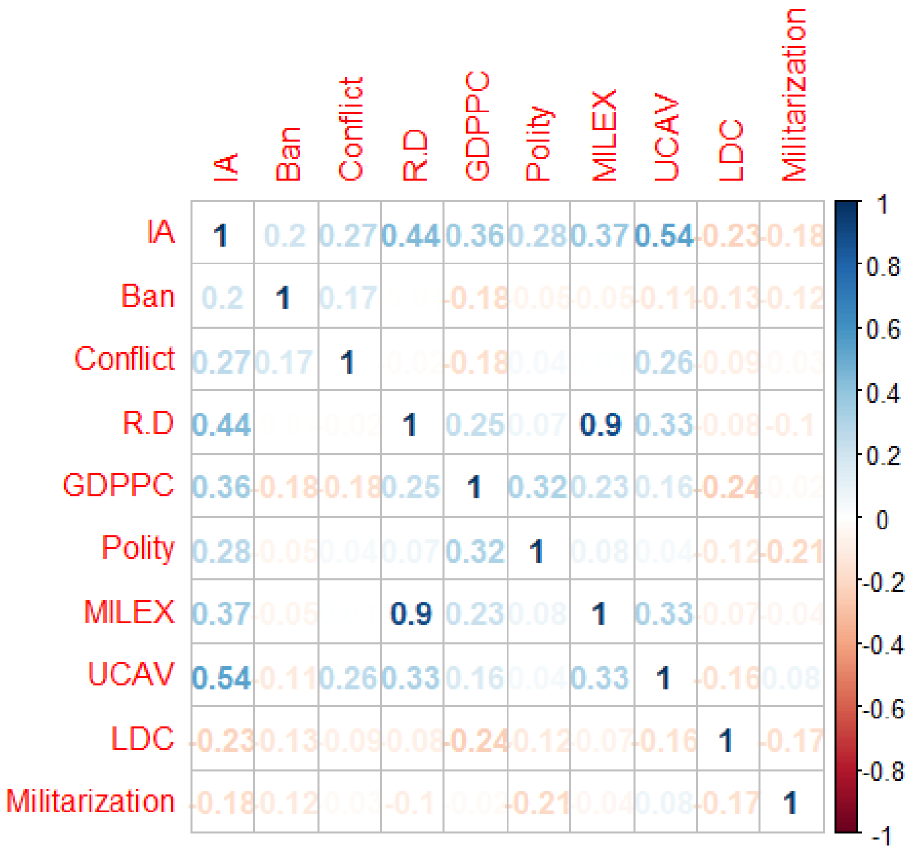
Figure 20: Distribution of Power across the Policy Continuum



Source: Author (2018)

Appendix 5: Correlation Matrix

Figure 22: Correlation Matrix



Source: Author (2018)

Charles University in Prague

Faculty of Social Sciences

Institute of Political Studies

Diploma Thesis Project

Campaign to stop 'killer robots': prospects of a preemptive ban on autonomous weapons systems

Name: Ondřej Rosendorf

Academic Year: 2017/2018

Date of submission: 30. 9. 2017

1 Introduction

In July 2015, the Future of Life Institute published an open letter on autonomous weapons which was calling for a ban on offensive autonomous weapons beyond meaningful human control. Since then, the letter has been signed by more than 3000 AI/Robotics researchers, and has gained much attention in the public (FLI, 2015). Above all, the letter contains a warning for national states wishing to pursue or already pursuing offensive AI capabilities: "[i]f any major military power pushes ahead with AI weapon development, a global arms race is virtually inevitable, and the endpoint of this technological trajectory is obvious: autonomous weapons will become the Kalashnikovs of tomorrow" (FLI, 2015).

Nevertheless, while the international movement to ban lethal autonomous weapons seems to be gaining momentum, especially with the onset of the Campaign to Stop "Killer Robots", there is a persistent and ongoing trend of increasing the autonomy in military systems of various nation states, including the U.S., U.K., China, Russia, Israel and others (see, Wareham and Goose, 2017; Kralingen, 2016: 136; Sharkey, 2012: 788). Today, governments around the globe recognize the benefits of autonomy[77]. This is highlighted, for example, by a recent quote by President Putin who said: "[w]hoever becomes the leader in this sphere [i.e. AI] will become the ruler of the world" (RT, 2017). Even though no state openly engages in AI arms race, some believe it's already too late to avoid it (Geist, 2016: 318). In turn, "[…] the continued failure of major military powers to agree to […] a ban on 'killer robots' suggests that one will not be reached […]" (Geist, 2016: 318). Critics argue that even if there was a ban, it is unlikely to be ratified globally, and violations are likely to occur at least to some extent (Kralingen, 2016: 150; Wallach, 2017: 29)[78].

To be sure, this rather pessimistic view is not shared by all academics. Amongst the most vocal proponents of the ban are the founders of the International Committee for Robots Arms Control (ICRAC) – Noel Sharkey and Peter Asaro (Marchant et al., 2011: 298). Peter Asaro, for example, criticized the idea expressed by Anderson and Waxman that a pre-emptive ban on lethal autonomy is unrealistic, ineffective, dangerous, and based on sci-fi scenarios of lethal autonomous weapons systems gone rogue (Anderson and Waxman, 2013: 3–4; Asaro, 2012: 707). According to Asaro, "[a]t the very least, a ban is not unrealistic in the sense that it might likely find broad public and official support" (Asaro, 2012: 707–708). While this may be true for most of the 94 states that attended the 2017 April meeting of the Convention on Certain Conventional Weapons (CCW) at the UN Office in Geneva

[77] The drivers behind military interest in increasingly autonomous technologies include, for example: force multiplications, expanding the battle-space, extending the warfighter's reach, casualty reduction, etc. (see, UNIDIR, 2014: 5–6; Marchant et al., 2011: 275).

[78] Wendell Wallach notes that even with a ban, "[t]he Great Powers, if they so desire, will find it easy to mask whether a weapon system has the capability of functioning autonomously" (Wallach, 2017: 34).

where informal discussions on a preemptive ban are held, to this day only 19 states have openly embraced this solution (Wallach, 2017: 28).

According to his account of the aforementioned meeting, Wendell Wallach noted that even though the large military powers such as the U.S. have expressed publicly their support for a ban under condition that such proposal would be supported by the majority of member states, "[b]ehind the scenes […] the principal powers express their serious disinclination to embrace a ban." He added, "[m]any of the smaller states will follow their lead" (Wallach, 2017: 28). It must be recognized that, while a pre-emptive ban on lethal autonomy may not be the most preferred approach to dealing with threats posed by lethal autonomy, virtually any form of regulation may be better than no regulation at all. A poorly managed AI arms race could lead to disastrous outcomes (Geist, 2016: 318). It is precisely this disastrous vision what makes autonomous weapons a compelling case for arms control. Regardless of whether we accept the notion of evitability/inevitability of an AI arms race, it is clear that some constraints need to be put in place.

This thesis aims to map the current discussions on lethal autonomy. In particular, it aims to identify the actors, namely national states, involved in the debate on the international level, their stances on ethical and operational implications of AWS use, their preferences in terms of possible arms control solutions, and an overall trajectory of the debate. The purpose of this endeavor is to identify policy outcomes that are likely to garner the most support among actors. Results of this research can then be used to predict behavior of individual states when presented with different policy outcomes. With a view to the above mentioned, this thesis seeks to provide an answer to the following research question.

Research question: Which one of the available policy solutions—ranging from the status quo to a preemptive ban—will likely gain the most support among actors involved in discussions on lethal autonomy?

By answering the above stated research question, the thesis will also provide an answer to one of the many related questions. For example, will there be enough support for the pre-emptive ban on autonomous weapons? Will there be enough support for a moratorium on the development of autonomous weapons? Will the majority of states resort to national self-regulation? Or, will most of the states retain the wait-and-see posture? The thesis employs a game-theoretical approach, in particular, the median voter theorem to provide an answer to these questions. In addition, the thesis will derive key concepts and theoretical assumptions from theories of arms control.

2 Literature review

Purpose of this literature review is, first and foremost, to identify available policy outcomes in regards to arms control of autonomous weapons systems. At the same time, it is important to note that this thesis does not seek to evaluate legal implications of autonomous weapons use as such. This section merely seeks to establish what options are available to decision-makers. The range of policy options will be later operationalized into a policy continuum for the purposes of game-theoretical analysis. Second, the literature review will identify key actors involved in the debate. These include, in particular, major military powers, countries developing offensive AI capabilities, and, potentially, international organizations and NGOs.

In their review of potential legal and policy constraints on autonomous weapons, Marchant et al. (2011) distinguish between soft and hard law/governance approaches. First, soft law approach includes options such as codes of conduct, transgovernmental dialogue, information sharing and CBMs, or framework conventions. Second, hard law approach includes options such as prohibitions and restrictions related to acquisition, research and development, testing, deployment, transfer or proliferation, and use. In addition, there is also an option of no change, whereby principles of LOAC alone will provide a mode of governance over autonomous weapons[79] (Marchant et al., 2011: 295–314).

Although all of the above mentioned options are on the table, the authors note that a more formal and traditional approach, i.e. binding international arms control agreement, is more likely as with other new weapons categories (Marchant et al., 2011: 298). Specific examples of legal and policy options aimed at constraining autonomous weapons systems include, for example, a ban, a moratorium on development, but also national self-regulation of development and deployment of autonomous weapons systems, or a posture of wait-and-see (UNIDIT, 2014: 1). Some argue for the incorporation of certain measures into already existing multilateral legal frameworks such as the CCW (Marchant et al., 2011: 299–300)[80].

As regards the second purpose of the literature review, it is important to recognize the unique position of the U.S., as both a super power and the leader in developing autonomous weapons systems. It is believed that the U.S. has the power to "[…] establish the precedents and norms that will shape the future of armed conflict" (Asaro, 2012: 707).

[79] Principles of LOAC include: military necessity, proportionality, discrimination, command responsibility, and general prohibition on the employment of weapons of a nature to cause superfluous injury or unnecessary suffering (Marchant et al., 2011: 295).

[80] One of the reasons why recent efforts at the UN, CCW, have failed is the inability of major military powers to agree on a shared definition of an autonomous weapon (Geist, 2016: 319). It should be mentioned that this issue is likely to come up in discussions about any form of arms control, and is not specifically limited to a preemptive ban (see, Marchant et al., 2011: 290; UNIDIR, 2014: 3–4).

However, as suggested by Wendell Wallach, behind the scenes, major powers such as the U.S. has expressed reluctance to embrace a ban (Wallach, 2017: 28). A ban on autonomous weapons is unlikely to be ratified by any major power unless other major powers cooperate, since this would mean a serious disadvantage in combat—something that no major power would allow to happen (Kralingen, 2016: 150).

Of the total 81 countries that have publicly elaborated their views on lethal autonomous weapons, only 19 countries have explicitly called for a ban. These include: Algeria, Argentina, Bolivia, Chile, Costa Rica, Cuba, Ecuador, Egypt, Ghana, Guatemala, Holy See, Mexico, Nicaragua, Pakistan, Panama, Peru, State of Palestine, Venezuela, and Zimbabwe (CSKR, 2017: 1). Perhaps on the opposite side discussions are the top five weapons exporting countries that have developed and deployed weapon systems with varying degree of autonomy. These are: the U.S., Russia, China, France, and Germany (Ricks, 2016)[81]. Other major powers such as the U.K. have opposed a preemptive ban in the past (Bowcott, 2015).

3 Theoretical framework

As was mentioned earlier, this thesis draws key concepts and theoretical assumptions from theories of arms control, and, by extension, contingent realism. This variant of realism is particularly useful because, unlike structural realism, contingent realism, which is associated with Charles L. Glaser, does not take a purely zero-sum view of international politics. Therefore, cooperation can be a preferred option for states under certain conditions. Similarly to other variants of structural realism, contingent realism builds on the assumption of rational state-actors in an anarchic system characterized by self-help strategies. However, there are certain areas where states may share a common interest (Selim, 2013: 11). One such common interest is to avoid arms race in the pursuit of military superiority—an issue which arms control seeks to address (Larsen, 2002: 2).

Arms control, according to Jeffrey A. Larsen, "[…] can be defined as any agreement among states to regulate some aspect of their military capability or potential. The agreement may apply to the location, amount, readiness, and types of military forces, weapons, and facilities" (Larsen, 2002: 1). In general, the main objectives of such agreements are to reduce the likelihood of war, reduce the political and economic costs of preparing for war, and minimize the scope and violence of war if it occurs (Schelling and Halperin, 1985: 3). Importantly, arms control is perceived as way of achieving a stable balance of power among actors (Larsen, 2002: 5). But, perhaps even more importantly, the

[81] Other countries in the dataset, which covers 284 weapons systems, include: Austria, Canada, India, Iran, Israel, Italy, Netherlands, South Korea, Sweden, Switzerland, or U.K. (Ricks, 2016). However, it is important to note that this list is by no means exhaustive.

initial reason for arms control was to enhance national security. In words of Headley Bull, arms control becomes a means to an end—i.e., enhancement of security (Bull, 1983: 21)[82]. Herein lies the rationale for self-interested state-actors to engage in cooperation. They do so in order to ensure their survival at the backdrop of a looming autonomous weapons systems arms race[83].

In addition to the aforementioned contingent realism and arms control theories, the thesis seeks to explore the role of stigmatization in arms control. Similarly to various weapon systems that have already been banned without requiring an inspection regime, for example, blinding lasers, prohibition of autonomous weapons will also rely on representatives of civilian society, NGOs in particular, to monitor and stigmatize violations (Wallach, 2017: 29)[84]. Nevertheless, according to Cooper (2011), "[…] the long history of pariah weapons regulation illustrates the way that weapons taboos frequently reflect the interests of the powerful" (Cooper, 2011: 143). Thus, there seems to be a link between the material and political interests of states on one hand, and the emergence of regulation on the other (Cooper, 2011: 143). Accordingly, actors interested in arms control regulation should prefer stigmatizing strategies.

Hypothesis 1 "looming arms race": Compared to status quo, all form of arms control regulation will, in total, garner more support among actors involved in discussions on lethal autonomy.

Hypothesis 2 "balance of power": Actors with limited autonomous weapons capabilities are more likely to support more restrictive forms of arms control regulation.

Hypothesis 3 "stigmatizing strategies": Actors preferring stigmatizing strategies are likely to support more restrictive arms control regulations.

4 Research design

[82] In some cases, arms control can lead to an increase in military forces, weapons, or facilities if such increase leads to an improved stability (Larsen, 2002: 3).

[83] In this context, Geist (2016) suggests that difficult compromises are inevitable for major military powers to achieve mutual security. However, besides maintaining regional and global stability, "[t]he major powers also share a common interest in limiting the ability of rogue states and terrorist groups to acquire and use military AI technologies" (Geist, 2016: 320).

[84] The crucial difference between autonomous weapons and various other weapon systems is, however, that lethal autonomy is a feature rather than a weapon system. Virtually any military system can have this feature, and it can be masked (Wallach, 2017: 29). Therefore, traditional verification regimes would fail to achieve the desired level of transparency to provide reassurance for actors involved.

In order to achieve the aforementioned aims, the thesis employs Q methodology. First, an extensive literature review will be conducted to identify the existing policy options—forms of arms control regulations—related to autonomous weapons. In this regard, a policy continuum will be created in order to capture available options in range of a pre-emptive ban and no regulation at all, i.e. status quo. Next, the author will identify the most important actors involved in the debate. Due to the scope of this research only a handful of key actors will be chosen. Actors with the same ordering of policy preferences may be grouped into blocs[85].

Second, the author will identify order of preferences, including intensity of preference, as regards the above mentioned policy options of individual actors. The author will utilize a literature review and a content analysis of messages derived from statements of key decision-makers, as well as available strategic documents, policy memos and official positions of the respective actors. This will be achieved by identifying specific keywords, e.g. "ban", prior to the execution of the analysis. Results will be summarized in frequency tables and graphs. Supplementary materials such as media reports may also be used.

Third, in addition to the above mentioned tables, the analysis will be accompanied by word clouds as a means of visual representation of data[86]. This will allow the author to identify some of the key themes in actors' statements on autonomous weapons, and to establish a broader picture of the debate that could be otherwise skewed by selected searching criteria. This step will help to establish the overall trajectory of the debate, identifying key themes that emerge in discussions.

Fourth, the median voter theorem will be employed to find out which policy outcome is likely to garner the most support among actors[87]. Assumptions of single-peaked preferences, the majority rule, and unidimensionality will be addressed in this process. Furthermore, the variable of power will be added to individual actors. Finally, the thesis will explore another dimension of the debate by adding the dimension of stigmatization. In turn, a spatial model of the debate on lethal autonomy will be provided[88]. This will allow the author to establish, whether a compromise is possible between actors with different arms control policy preferences while controlling for stigmatization strategies. This issue-linkage is

[85] This could include, e.g. the aforementioned 19 states that have explicitly called for a ban on autonomous weapons systems (see, CSKR, 2017).

[86] Word clouds are "[...] a method of visually representing text data [...] typically used as a summative way to display the frequency of keywords [...]" or "[...] providing a visual of the most important topics [...]" (Ramlo, 2011: 100–101).

[87] Put simply, "[t]he median voter theorem says that, if issues are one-dimensional, preferences are single-peaked, and it takes a majority to win, then the median voter's position is the winning outcome" (Bueno de Mesquita 2010: 48).

[88] "Spatial models are a class of abstract perspectives that assume that we can locate decision makers and their policy preferences either on a line or continuum or in a [conceptual] space that includes more than one dimension" (Bueno de Mesquita 2010: 47).

conducted, partially, to find out whether there are any alternative win-sets between actors that may otherwise lose to the unidimensional median voter.

5 References

Anderson, K. – Waxman, M. (2013). Law and Ethics for Autonomous Weapon Systems: Why a Ban Won't Work and How the Laws of War Can. *Hoover Institution, Research Paper* (available at: http://www.hoover.org/research/law-and-ethics-autonomous-weapon-systems-why-ban-wont-work-and-how-laws-war-can, 25. 9. 2017).

Asaro, P. (2012). On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making. *International Review of the Red Cross* 94 (886), pp. 687–709.

Bowcott, O. (2015). UK opposes international ban on developing 'killer robots'. *The Guardian* (available at: https://www.theguardian.com/politics/2015/apr/13/uk-opposes-international-ban-on-developing-killer-robots, 27. 9. 2017)

Bueno de Mesquita, B. (2010). *Principles of international politics, 4th edition* (Washington, DC: CQ Press).

Bull, H. (1983). The Traditional Approach to Arms Control Twenty Years After. In: Nerlich, U. ed., *Soviet Power and Western Negotiating Policies* (Cambridge, MA: Ballinger), pp. 21–30.

Cooper, N. (2011). Humanitarian Arms Control and Processes of Securitization: Moving Weapons along the Security Continuum. *Contemporary Policy* 32 (1), pp. 134–158.

CSKR (2017). Country Views on Killer Robots, 23 May 2017. *Campaign to Stop Killer Robots* (available at: http://www.stopkillerrobots.org/wp-content/uploads/2013/03/KRC_CountryViews_May2017.pdf, 27. 9. 2017).

FLI (2015). *Autonomous Weapons: An Open Letter From AI & Robotics Researchers* (available at: https://futureoflife.org/open-letter-autonomous-weapons/, 27. 9. 2017).

Geist, E. (2016). It's already too late to stop the AI arms race—We must manage it instead. *Bulletin of the Atomic Scientists* 72 (5), pp. 318–321.

Goose, S. – Wareham, M. (2017). *The Growing International Movement Against Killer Robots* (available at: https://www.hrw.org/news/2017/01/05/growing-international-movement-against-killer-robots, 27. 9. 2017).

Kralingen, M. (2016). Use of Weapons: Should We Ban the Development of Autonomous Weapons Systems? *The International Journal of Intelligence, Security, and Public Affairs* 18 (2), pp. 132–156.

Larsen, J. (2002). An Introduction to Arms Control. In: Larsen, J. ed., *Arms Control: Cooperative Security in a Changing Environment* (Boulder, CO: Lynne Rienner Publishers), pp. 1–15.

Marchant, G. – Allenby, B. – Arkin, R. – Barrett, E. – Borenstein, J. – Gaudet, L. – Kittrie, O. – Lin, P. – Lucas, G. – O'Meara, R. – Silberman, J. (2011). International Governance of Autonomous Military Robots. *The Columbia Science and Technology Law Review* 12 (1), pp. 272–315.

Ramlo, S. (2011). Using word clouds to present Q methodology data and findings. *Human Subjectivity* 9 (2), pp. 99–111.

Ricks, T. (2016). Weapons autonomy is rocketing. *Foreign Policy* (available at: http://foreignpolicy.com/2016/09/28/weapons-autonomy-is-rocketing/, 27. 9. 2017).

RT (2017). *'Whoever leads in AI will rule the world': Putin to Russian children on Knowledge Day* (available at: https://www.rt.com/news/401731-ai-rule-world-putin/, 27. 9. 2017).

Schelling, T. – Halperin, M. (1985). *Strategy and Arms Control* (Washington, DC: Pergamon-Brassey's).

Selim, G. (2013). Arms Control and Security Cooperation: Contending Approaches. In: Selim, G., *Global and Regional Approaches to Arms Control in the Middle East* (Berlin: Springer), pp. 7–30.

Sharkey, N. (2012). The evitability of autonomous robot warfare. *International Review of the Red Cross* 94 (886), pp. 787–799.

UNIDIR (2014). *Framing Discussions on the Weaponization of Increasingly Autonomous Technologies* (available at: http://www.unidir.org/en/publications/framing-discussions-on-the-weaponization-of-increasingly-autonomous-technologies, 25. 9. 2017).

Wallach, W. (2017). Toward a Ban on Lethal Autonomous Weapons: Surmounting the Obstacles. *Communications of the ACM* 60 (6), pp. 28–34.

Wareham, M. – Goose, S. (2017). The Growing International Movement Against Killer Robots. *Human Rights Watch* (available at: https://www.hrw.org/news/2017/01/05/growing-international-movement-against-killer-robots, 25. 9. 2017).

6 Additional literature

Arkin, R. (2010). The Case for Ethical Autonomy in Unmanned Systems. *Journal of Military Ethics* 9 (4), pp. 332–341.

Boulanin, V. (2016). Mapping the Development of Autonomy in Weapon Systems: A primer on autonomy. *SIPRI, working paper* (available at: https://www.sipri.org/sites/default/files/Mapping-development-autonomy-in-weapon-systems.pdf, 25. 9. 2017).

Bull, H. (1976). Arms Control and World Order. *International Security* 1 (1), pp. 3–16.

Crootof, R. (2015). *The Killer Robots Are Here: Legal and Policy Implications* (available at: https://ssrn.com/abstract=2534567, 25. 9. 2017).

Glaser, C. (1995). Realists as optimists: Cooperation and self-help. *International Security* 19 (3), pp. 50–90.

Hoffman, R. – Cullen, T. – Hawley, J. (2016). The myths and costs of autonomous weapon systems. *Bulletin of the Atomic Scientists* 74 (4), pp. 247–255.

Jervis, R. (1993). Arms Control, Stability, and Causes of War. *Political Science Quarterly* 108 (2), pp. 239–253.

Lin, P. (2010). Ethical Blowback from Emerging Technologies. *Journal of Military Ethics* 9 (4), pp. 313–331.

Roberts, B. (1997). Arms control in the emerging strategic environment. *Contemporary Security Policy* 18 (1), pp. 57–82.

Roff, H. (2014). The Strategic Robot Problem: Lethal Autonomous Weapons in War. *Journal of Military Ethics* 13 (3), pp. 211–227.

Scharre, P. (2015). The Opportunity and Challenge of Autonomous Systems. In: Williams, A. – Scharre, P. eds., *Autonomous Systems: Issues for Defence Policymakers* (Hague: NATO Communications and Information Agency), pp. 3–26.

Sharkey, N. (2014). Towards a principle for the human supervisory control of robot weapons. *Politica & Società* 2, pp. 305–324.

Sharkey, N. (2012). Killing made easy: from joysticks to politics. In: Lin, P. – Abney, K. – Bekey, G. eds., *Robot ethics: the ethical and social implications of robotics* (Cambridge: MIT Press), pp. 111–128.

UNDIR (2015). *The Weaponization of Increasingly Autonomous Technologies: Considering Ethics and Social Values* (available at: http://www.unidir.org/files/publications/pdfs/considering-ethics-and-social-values-en-624.pdf, 25. 9. 2017).