SYM-AM-19-087

# PROCEEDINGS
## OF THE
# SIXTEENTH ANNUAL ACQUISITION RESEARCH SYMPOSIUM

## THURSDAY SESSIONS
## VOLUME II

**Acquisition Research:**
**Creating Synergy for Informed Change**

**May 8–9, 2019**

**Published: April 30, 2019**

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL

# The Effects of Exporting on Defense Acquisition Outcomes: A Quantitative Look at FMS Contracting—Preliminary Findings

**Andrew Hunter—**is a senior fellow in the International Security Program and director of the Defense-Industrial Initiatives Group at CSIS. From 2011 to 2014, he served as a senior executive in the Department of Defense, serving first as Chief of Staff to Under Secretaries of Defense (AT&L) Ashton B. Carter and Frank Kendall, before directing the Joint Rapid Acquisition Cell. From 2005 to 2011, Hunter served as a professional staff member of the House Armed Services Committee. He holds an MA in applied economics from Johns Hopkins University and a BA in social studies from Harvard University.

**Samantha Cohen—**worked as a research associate with the Defense-Industrial Initiatives Group at CSIS. Her work focuses on managing and analyzing data to identify relationships among policies, defense spending, and the related impacts on the United States and national security. Her recent research focused on the design and management of international joint development programs, new entrants' survival rates and business graduation in the market for government contracts, and defense acquisition trends. Cohen holds a BS in economics from American University in Washington, DC, and an MS in economics from Katholieke Universiteit (KU) in Leuven, Belgium.

**Greg Sanders—**is a fellow in the International Security Program and deputy director of the Defense-Industrial Initiatives Group at CSIS, where he manages a research team that analyzes data on U.S. government contract spending and other budget and acquisition issues. In support of these goals, he employs SQL Server, as well as the statistical programming language R. Sanders holds an MA in international studies from the University of Denver and a BA in government and politics, as well as a BS in computer science, from the University of Maryland. [gsanders@csis.org]

**Nicholas Halterman—**is an intern with the Defense-Industrial Initiatives Group at CSIS. His current work focuses on statistical analysis and machine learning in Python and R. He holds a BA in politics, philosophy, and economics from Oxford University, where he focused on quantitative political methods and political theory. Halterman's previous research was on civilian displacement and charitable foundation behavior. [nhalterman@csis.org]

## Contributors:

**Maggie X. Wang—**was an intern with the Defense-Industrial Initiatives Group (DIIG) at CSIS. In addition to her work at DIIG, which focused on arms transfers data, she has conducted open source research for non-proliferation organizations and political risk consultancies. She is pursuing a BA in government and data science at Smith College and is currently studying at the Graduate Institute of International and Development Studies in Geneva.

**Andrew Howe—**was an intern with the Defense-Industrial Initiatives Group (DIIG) at CSIS. In addition to his work with international military transfers data, he also works on developing interactive data visualizations. He holds a BA in human evolutionary biology from Harvard University and an MPP from Georgetown University's McCourt School of Public Policy.

## Abstract

This paper studies how approaches to security cooperation as well as the characteristics of foreign military sales (FMS) recipients influence defense acquisition outcomes. A review of the literature finds that the lower asset specificity for internationally traded goods, the strength and history of the security relationship, and the quality of partner institutions all are likely influencers of performance. This project has created a unique contract-level FMS dataset, cross-referenced other sources to evaluate the quality of contract reporting, and used to validate economic research regarding the influence of international sales transaction cost.

## Introduction

Foreign military sales (FMS) have grown markedly in recent years with major agreements announced during the prior administration, followed by a broad-based push to accelerate and increase FMS by the present administration, which includes revisions to the Conventional Arms Transfer Policy. This trend, shown in

, is more prominent in agreements than deliveries, although the latter have been increasing and jumped in 2017. This FMS drive has multiple sources, not the least of which being a greater emphasis on working by, with, and through partner nations in the National Defense Strategy potentially overlapping with the economic challenges of the global financial crisis and subsequent U.S. spending reductions that reduced defense industry revenues. These twinned motivations are important because arms exports, as recognized by U.S. law, are political and a form of security cooperation, while at the same time having economic and industrial base implications. The political challenges of arms exports have been thrown into sharp relief by the ongoing debate over U.S. support for the Saudi led war in Yemen, as opponents of the war have sought to cut back FMS as a way of adding to the pressure they seek to apply to the Saudi regime.

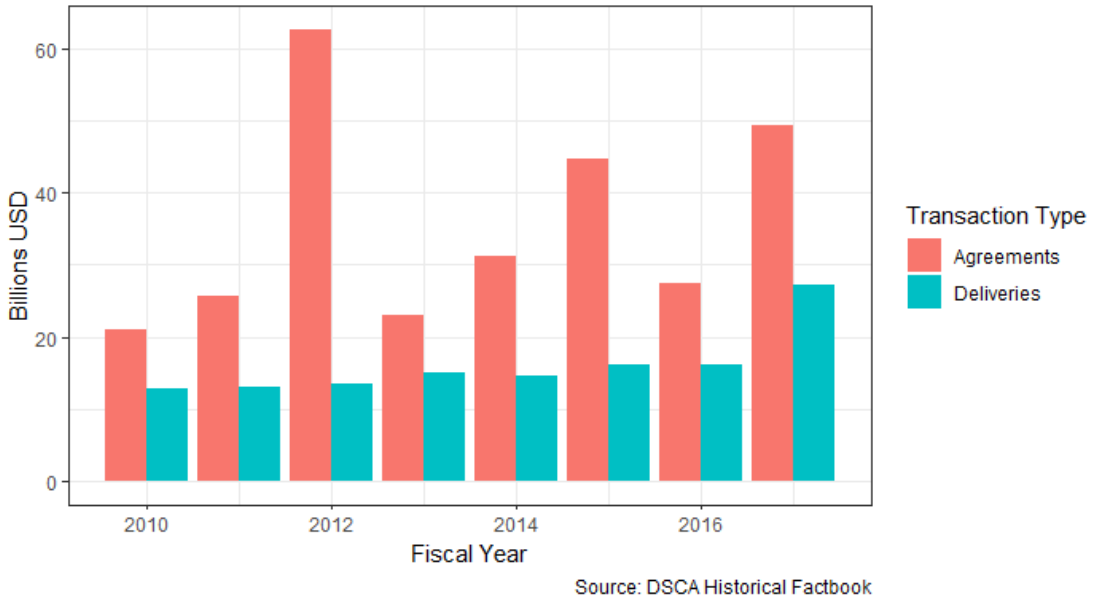


**Figure 1. FMS Agreements and Deliveries by Fiscal Year**

Given FMS's utilization of the U.S. defense acquisition system, and in keeping with laws emphasizing foreign policy considerations in all arms exports, those emphasizing economic and industrial base factors tend to also posit that expanding FMS furthers broad U.S. national security goals. Likewise, those emphasizing deliberation and caution point to the risks of poorly considered deals falling apart, and of the possible proliferation of closely held U.S. technological developments, potentially undermining U.S. national security goals.

The interaction of these considerations means that when considering the acquisition effects of FMS specifically, a wide range of potential influences come into play. On the one hand, the effects of sequestration have incentivized industry to look abroad for revenue growth, and program managers have looked to capitalize on budget savings from overseas

sales that can result in lower production costs and shared support costs. On the other hand, arms exports are inherently challenging due to the risk of complications when meeting foreign requirements, instability in international demand, blocks by Congress or the executive branch, and the risk of adverse technology proliferation. This project seeks to evaluate the performance of federal contracts that incorporate FMS. This paper presents the work done to identify the appropriate literature and hypotheses and to build a curated dataset of federal contracts that utilize FMS.

To evaluate the performance of contracts that utilize FMS, the study team first references the existing body of literature that analyzes contract performance and investigates if any papers specifically looked at contract performance for FMS. While the body of contract performance literature is extensive, there are no pieces that empirically analyzed contract performance for FMS contracts. This is likely because the publicly available contracting data from the Federal Procurement Data System (FPDS) is incomplete in indicating whether a transaction was FMS. Thus, a large portion of work done for this paper involves curating a dataset using other fields in FPDS to identify FMS transactions that were mislabeled or unlabeled. This labeling effort includes both application of rules based on transaction funding account and an experiment with labeling using machine learning detailed later in this paper.

While previous work has not examined FMS contracts in particular, the existing bodies of literature provide guidance on theorizing about and measuring contracting performance. Work on security assistance details some inherent challenges of arms exports in meeting foreign requirements and the risk of adverse technology proliferation. Existing work on transaction cost theory provides a foundation to build models that estimate the effects of FMS contract characteristics on FMS contract performance outcomes. Several authors have found that transaction costs, and in particular asset specificity, are a driving force behind acquisition outcomes for services and products (Williamson, 1981; Brown & Potoski, 2003; Adler et al., 1998). Expansion to international markets may reduce asset specificity, as well as creating other economies of scale. Other research, however, has examined how transaction costs are exacerbated in the context of international business (Berghuis & Butter, 2017). This paper will draw on these theories to explore contract performance in the context of FMS.

### *Scope*

To guide the research for this project, the study team posed these four research questions, the first two of which are answered in this paper:

1. How can contracts that utilize FMS be better identified in FPDS using information from other fields?
2. How does FPDS foreign funding data align with the Defense Security Cooperation Agency's FMS data?
3. Do FMS contracts perform better than non-FMS contracts? This question was subsequently expanded to cover projects incorporating FMS and not just FMS transactions.
4. What variables contribute to the performance of FMS contracts and in what direction and magnitude?

The remainder of this paper expands on the issues raised above by discussing the FMS process, its important role in the defense industrial base, and how contract theory informs the analysis for this project. It also details the methodology used to identify FMS

contracts in FPDS, the resulting database, and provides a descriptive overview of the distribution of contract performance metrics for FMS and non-FMS contracts.

## Literature Review

### *Security Assistance and Cooperation*

FMS is intended as a U.S. foreign policy tool for strengthening the security of the United States and promoting global security. FMS is authorized under Section 3 of the Arms Export Control Act (AECA) where it is considered as security assistance. The *DoD Security Assistance Management Manual* (DoD 5105.38-M) has a list of eligible countries and organizations who can participate in FMS. An FMS process begins when a foreign customer determines that its military and security needs require a U.S. defense article or service. That foreign government or organization then alerts the U.S. government of its intent to participate in FMS through submitting a letter of request (LOR). From there, the U.S. government organization that is both relevant to the requirement and authorized to receive and process LORs, otherwise known as the implementing agency, works through an interagency process to determine whether the LOR requestor is an eligible participant of the security assistance process under AECA. If so, the implementing agency moves forward in determining an appropriate letter of offer and acceptance (Defense Security Cooperation Agency, 2012, C5.1).

### *Export Controls and End Use Monitoring*

Monitoring and evaluation are essential components of any form of security sector assistance. Throughout security assistance relationships, the United States is able to calculate return on investment, identify and prevent abuse of military resources, and enforce any forms of conditionality on assistance (Dalton et al., 2018, p. 9). In addition to its strategic importance, monitoring is statutorily required under the Leahy laws, which mandate vetting of individuals and units before they receive training or equipment, thereby preventing security sector assistance from going to foreign security forces that commit gross violations of human rights. Beyond the Leahy requirement for end-use monitoring, the Arms Export Control Act (AECA) and International Traffic in Arms Regulations (ITAR) place substantial restrictions and requirements on both foreign military sales and direct commercial sales, including requirements about the eligibility of potential recipient countries and eligibility of platforms and technologies (Gilman, 2014, p. 4). Two separate programs exist to provide end-use monitoring for transfers of military equipment: Blue Lantern and Golden Sentry. Blue Lantern operates under the Department of State's Directorate of the Defense Trade Controls and monitors use of equipment from direct commercial sales, while the Golden Sentry program is administered by the Defense Security Cooperation Agency and monitors FMS (Fergusson & Kerr, 2017, p. 6). Golden Sentry provides oversight for recipient security and handling of materials, reports any misuse or illegal transfer of equipment, and performs physical inspections and inventory management in some cases (Little, 2017).

Golden Sentry and other end-use monitoring are essential to reducing the risks of transfers by "ensuring that they are not misused and remain within the security force to which they are assigned" (Dalton et al., 2018, p. 10). Alongside concerns about human rights violations and potential proliferation of weapons beyond intended recipients, FMS can increase the risk of harmful strategic behavior by recipients. Capability transfers and the perception of U.S. support create moral hazards for recipient regimes, leading to opportunistic behavior like coup-proofing and power consolidation, both of which can ultimately degrade military capacity and undermine U.S. goals in security assistance (Boutton, 2018, pp. 8–10). These risks, and the monitoring needed to mitigate them, can significantly complicate security assistance and impose meaningful transaction costs.

Despite an increasing emphasis on the economic aspects of FMS in current political discourse, it remains the case that "arms transfers are a foreign policy tool and cannot be wholly separated from U.S. security cooperation policy" (Dalton, 2018, p. 38).

### *Defense Institutional Capacity*

Defense Institution Building (DIB) is an element of security cooperation which has received increased attention in recent years. It seeks to improve security outcomes and mitigate risk of material misuse by increasing institutional capacity in recipient countries to combat the dangers of instability, weak oversight, and poor governance (Dalton et al., 2018, p. 19). DIB is stipulated as an integral part of any security cooperation agreement, as part of the FY 2016 NDAA. The growing focus on DIB and on recipient-country institutions more broadly highlights the fundamentally political aspect of successful security assistance, including FMS. While FMS programs may not themselves involve significant DIB activities, the presence (or lack) of institutional capacity in recipient countries remains a central driver of risk.

### *Interoperability*

The 2018 National Defense Strategy expressed a clear desire to increase interoperability, noting that the ability to "act together coherently and effectively to achieve military objectives requires interoperability" (p. 9). While interoperability includes elements of communication and operational concepts, material overlap between forces can also be a significant contributor to interoperability. As De Vore (2011) argues, "States equipped with the same weapons can support, reinforce, repair, and resupply each other's armed forces without advanced warning" (p. 628). Combined with the shared training and logistic integration that can accompany arms transfers, FMS can provide the material foundation for increased interoperability between U.S. forces and recipient-nation forces. This line of reasoning is echoed in the 2018 National Defense Strategy, which includes as part of its plan for increasing interoperability the need to "prioritize requests for U.S. military equipment sales" (p. 9).

## *Economics for International Cooperation*

The rise in foreign military sales observed at the beginning of this report has been driven not just by security assistance concerns, but also economic factors. The Great Recession put pressure on defense budgets in the United States and Europe while expenditures increased for "several countries—particularly in East Asia, South Asia, the Middle East, and South America" (Gilman, 2014, p. 1). The present U.S. administration prominently featured economic ends in the April 19, 2018 National Security Presidential Memorandum Regarding U.S. Conventional Arms Transfer (CAT) Policy. That document made it a policy of the executive branch to

> increase trade opportunities for United States companies, including by supporting United States industry with appropriate advocacy and trade promotion activities and by simplifying the United States regulatory environment; strengthen the manufacturing and defense industrial base and lower unit costs for the United States and our allies and partners, including by improving financing options and increasing contract flexibility; facilitate ally and partner efforts, through United States sales and security cooperation efforts, to reduce the risk of national or coalition operations causing civilian harm. (Trump, 2018)

At the announcement briefing Dr. Peter Navarro, Assistant to the President for Trade and Manufacturing Policy, discussed these rationales and, when asked about the desire by

some buyers for offsets and technology transfer, went further to make the case for jobs and industrial promotion saying, “The organizational culture of the Trump administration is: buy American, hire American” (U.S. Department of State, 2018). With regards to the U.S. industrial base, the most explicit discussion of how economics and industrial issues tie into larger U.S. defense goals is the 2018 Report to the President by the Interagency Task Force in Fulfillment of Executive Order 13805, otherwise known as the Defense Industrial Base Review (IBR). This document points to concerns that prominently feature the first and second order effects from the Budget Control Act of 2011 (BCA) and sequestration (which will be referred to as the defense drawdown henceforth), which helped prompt a greater emphasis on foreign military sales.

### *The Defense Industrial Base Review*

A combination of the 2008 financial crisis, 2011 debt-ceiling crisis, 2011 closing of the Iraq War, and BCA led the domestic demand for defense items to decline. Specifically, the budget caps mandated by the BCA from fiscal year (FY) 2012–2021 were significantly lower than requested funding levels which triggered sequestration in 2013. A previous CSIS study found that the decline in budget carried over to the defense industrial base, which experienced decreased revenue across all platform portfolios:

> CSIS analysis showed that buried within the substantial decline in defense contract obligations were significant variation from sector to sector, with declines varying from catastrophic (Land Vehicles), to steep (Facilities and Construction, Space Systems), to relatively modest (Ships & Submarines). Other sectors suffered a whipsaw effect in which solid business growth suddenly switched to sharp decline (Aircraft). (McCormick, Hunter, & Sanders, 2017, p. VI)

Moreover, medium and large federal vendors experienced the most variance in defense market share and the top companies working with the DoD saw their portfolios shift from R&D to products and services (McCormick et al., 2018). The IBR (2018) also found that sequestration has led to lower defense spending compared to the levels projected before it was triggered.

The IBR has deemed sequestration as one of the five macro forces behind the risks that threaten the U.S. industrial base. The IBR discusses multiple ways in which sequestration causes risks to the industrial base: “inconsistent appropriations, uncertainty about future budgets, macro-level ambiguity in U.S. Government expenditures, and the effects of the Budget Control Act” (IBR, 2018). The IBR argues that successful markets are dependent on predictability, where industries can invest and plan based on informed decisions. That said, Harrison and Daniels (2018) note that while the budget caps drove a gap between Obama administration budget proposals and actual results, the challenges in relying on the DoD’s Five Year Defense Plan (FYDP) long predated the BCA:

> While the FYDP is useful for planning purposes, in the past, it has been a poor indicator of where the budget is headed. As shown in Figure 2, the FYDPs submitted by the Reagan administration greatly exceeded the actual level of funding appropriated by Congress, and the Reagan FYDPs continued projecting growth even when the budget was declining. In the 1990s, the Clinton administration repeatedly projected a lower defense budget than Congress ultimately appropriated. (Harrison & Daniels, 2018, p. 4)

The challenges of predicting did not go away even during the period of single party control of the Congress and the Executive Branch during the 114th Congress. As Harrison and Daniels (2018) note, “While the NDS calls for a ‘more resource-sustainable approach’ to fund this modernization effort, the unclassified summary of the strategy fails to delineate how it plans to fund its ambitions” (p. 1).

However, all aspects of the present difficulty in predicting the demand for defense goods and services are familiar from prior eras. As noted by the Interagency Task Force’s IBR, the spending uncertainty caused by sequestration often results in “peaks of surge and valleys of drought,” that disrupt scale production because suppliers can be left with excess capacity during the valleys of drought (p. 21). This can lead to long-term market distortion.

Lastly, the fluctuations in demand caused by the BCA have had rippling effects across defense industry supply chains where companies have struggled in their abilities to hire and retain the necessary skilled workforces for their products and services. While McCormick found that the U.S. subcontracting data was inadequate to fully examine supply chain questions, he did find “the market shock of sequestration and the defense drawdown had a disproportionate effect on Small and Medium-sized vendors” (McCormick et al., 2018, p. 17). The IBR adds that, “Without correcting or mitigating this U.S. Government-inflicted damage, DoD will be increasingly challenged to ensure a secure and viable supply chain for the platforms critical to sustaining American military dominance” (p. 21).

### *Transaction Cost*

Transaction cost theory, as a general approach to understanding economic behavior, lays the foundations for analysis of contracts. As defined by Williamson (1981), transaction cost theory measures transaction costs along three dimensions: frequency, uncertainty, and asset specificity; with asset specificity especially relevant to defense contracting. Minimizing transaction costs is a main driver of municipal governments’ decisions to contract services or products, and the type of transaction cost specific to a product or service plays a role in determining contract type. Thus, they are a strong driver of contract design and behavior (Brown & Potoski, 2003; Adler et al., 1998). In the context of military sales, FMS may raise costs for specific transactions due to the difficulties of international transfers, but it may also reduce transaction costs for overall projects by increasing economies of scale and reducing asset specificity. These effects are discussed in turn below.

### *International Supply Chains*

Berghuis and Butter (2017) studied transaction costs in the context of international supply chains and found that international contracting has characteristics that result in high “intangible” transaction costs that require contracts that are more detailed, complete, difficult, expensive, and that need higher-trust relationships. A previous CSIS study found that international acquisition programs “exhibit a greater level of inherent organizational complexity, which poses a range of obstacles…international programs encourage participants to behave opportunistically, face collective tradeoffs that result in sub-optimal end products for individual nations, and experience competing factors within their structure” (Sanders & Cohen, 2017). The study also found cases where the desired benefits were outweighed by adverse effects of international cooperation resulting in negative cost, schedule, and end-product outcomes. Berghuis and Butter (2017) note that these effects vary greatly based on the strength of the relationship between international partners, raising the possibility of measures of “relational contracting” which may result in superior performance.

### *Offsets*

Offsets are a central and contentious aspect of international defense sales. Offsets are accompanying agreements to defense sales which require sellers to provide some economic value to the purchasing country as part of the terms of service. They may be direct, such as a requirement for the seller to purchase components from the buyer country, or indirect, such as a requirement for the seller to purchase or invest in goods or services unrelated to the military sale (Petty, 1999). For military sales conducted through FMS, federal policy is that the "DoD does not encourage, enter into, or commit U.S. firms to FMS offset arrangements" (Acquisitions for Foreign Military Sales, 48 U.S.C. §225.7306). This policy does not, however, prevent U.S. firms from negotiating offsets as part of an FMS sale without direct DoD involvement.

Offsets in international defense sales raise potential issues for domestic economic benefits. Offset agreements may shift economic gains from production to host countries via local co-production or components restrictions, reduce competitiveness through technology and capacity transfers, and ultimately reduce or outweigh some of the economic benefits of FMS (Petty, 1999). Recently, the DoD's stance on offsets in FMS has grown more supportive, including a reduction in oversight of offsets negotiated between contractors and foreign customers (Censer, 2018). Overall, both the transactional burden of negotiating offsets and the potential economic harms to U.S. production pose a theoretical challenge to the economic benefits of FMS.

### *Asset Specificity*

While both international transaction costs and offsets pose challenges to the benefits of FMS, one strong argument for its benefits is the potential effect on asset specificity. For most procurement contracts, producing the final product requires significant investment in capital infrastructure, both physical and informational. Asset specificity refers to the level of specialization for that infrastructure (Williamson, 1981). When infrastructure can be used after contract completion to produce products for the open market or other contracts, the effective cost of investment for the supplier is decreased. When the infrastructure is specific to the current contract, as is frequently the case in the defense industry, the full cost of investment is borne by the supplier for that specific contract. Capital useful for post-contract production is effectively subsidized by that future revenue, while fully-specific infrastructure is not. The degree of asset specificity is therefore a crucial determinant of both contract price and degree of supplier investment. Where asset-specificity is high, infrastructure investment by the supplier is costlier and is thereby disincentivized. This can lead to under-investment and sub-optimal contracts or require costly monitoring and incentives to ensure adequate investment (Schmitz, 2001).

FMS offers a potential boon to the United States government by reducing asset specificity. Asset specificity is high in defense contracting because it is typically a monopsony and requires highly specialized technical capacity, typically leaving suppliers with expensive infrastructure that cannot be reused after a defense contract expires. We should expect this to significantly increase prices: Defense contracts experience high costs to infrastructure investment and require significant incentives (and accompanying monitoring) to overcome those costs and achieve an optimal product (Schmitz, 2001). FMS, however, alleviates the effects of monopsony, and allows for potential asset-reuse after a U.S. government procurement contract expires. While the infrastructure remains specific to a technological product, it becomes viable for use in multiple contracts with multiple recipients. In short, the infrastructure may only be useful for producing F-15s, but producing F-15s for the United States, United Kingdom, and so forth, effectively reduces asset specificity by increasing the applications for the infrastructure.

Notably, this relies on the supplier *expecting* these future contracts. When firms know that FMS will occur, their estimates of asset-specificity should decrease, leading to increased investment and superior outcomes (Schmitz, 2001). This theory suggests that contracts including FMS from the outset with defense exportability features should have lower costs and superior outcomes to equivalent contracts that do not, and that the earlier in the process that FMS is included the stronger these effects should be.

### *Advantages of Scale*

Alongside asset specificity, increases in scale can improve acquisition outcomes through other mechanisms. While asset specificity helps improve outcomes by increasing incentives to invest in up-front capital and training, high production output can help reduce per-unit costs of investment and training. Holding up-front costs constant, each additional unit of production reduces the average per-unit cost until it approaches the marginal cost of each new unit. This economy of scale is central to the effects of monopolies, in which potential harms of market consolidation must be weighed against the benefits of decreasing per-unit cost with increasing scale (Peltzman, 1997). Alongside the declining per-unit cost of infrastructure, increased scale carries benefits through learning curves. As production occurs, involved workers gain experience and tend to discover more efficient techniques, leading to a declining *marginal* cost to production, on top of the declining *average* cost to production experienced for physical infrastructure (Sanders & Huitink, 2019). Unlike in the case of domestic monopolies, FMS does not clearly reduce domestic competition in order to achieve gains in scale, but effectively creates new customers by expanding the potential pool of buyers to foreign governments. This may allow FMS to achieve economies of scale for defense industrial producers without making the traditional tradeoffs to competition experienced in domestic situations.

Economic benefits from decreasing unit cost last beyond the time of purchase. When FMS and domestic procurement run concurrently, economies of scale and learning curve benefits can extend to maintenance, upgrades, and other lifecycle costs, particularly as many FMS products require additional service and parts from the United States after the initial sale. In general, Kirkpatrick (2004) finds that lower per-unit costs are associated with lower lifecycle costs, indicating not only a direct economy of scale to maintenance and parts, but a follow-on effect from reducing initial unit costs. Taken together, these effects offer a theoretical case for FMS lowering per-unit and lifecycle costs, both of which could drive superior acquisition outcomes for programs and platforms which include FMS.

Finally, FMS transactions do not only affect the immediate production cycle, but may have lasting effects on communication, infrastructure, and future projects. The IBR highlights the importance of maintaining and growing defense cooperation agreements with partners and allies to achieve economies of scale and scope as well as interoperability. Specifically, the IBR mentions the FY2017 NDAA's addition of Australia and the United Kingdom to the National Technology Industrial Base (NTIB) as an opportunity to jointly work on industrial base challenges (Interagency Task Force, 2018). The FMS process may help establish and grow defense cooperation by providing U.S. produced materials, ultimately creating the conditions for joint development, DCS, or other forms of security sector cooperation which may carry economic benefits for the U.S. defense industrial base.

## Conceptual Framework and Hypotheses

This paper posits that a range of considerations from the security cooperation and assistance domain, as well as traditional economics and contracting literature, have a relationship with foreign military sales contracting outcomes. On both the positive and negative sides of the ledger, strategic and political considerations by the United States and

the purchaser nation may influence the level of support for the program and whether it completes delivery at all. Transaction costs literature, organizational complexity, and traditional considerations of scale provide a possible mechanism for these non-economic considerations to influence outcomes as the purchase quantities, supporting institutional infrastructure, and alignment of U.S. and recipients' interests all depend on a variety of factors that can be better measured at the country-level rather than being specific to any given project.

Before testing these hypotheses, the study team had two falsifiable premises to test, which are the focus of this paper. These two premises directly relate to the study's research questions and must be confirmed before the study team can have confidence in the dataset produced as part of this project.

> $P_1$: Foreign Military Sales data identifiable in the Federal Procurement System correlates with and captures a majority of the spending reporting from other sources.

As will be discussed in subsequent sections, the official labeling of FMS contracts is radically incomplete prior to 2016. The results section includes comparisons of FPDS data with that of the Defense Security Cooperation Agency (DSCA) and the Stockholm International Peace Research Institute's (SIPRI) Arms Transfer database. While both are inexact comparisons, this cross validation is critical to establishing appropriate level of confidence and caveats for use of the FPDS dataset.

> $P_2$: The pricing mechanism selected for FMS contracts will vary significantly from comparable domestic contracts.

The transaction cost literature emphasizes that acquirers respond to different transaction cost context with different forms of contracts. This observed property provides a useful way to validate the relevant of transaction costs considerations. In addition, even when an FPDS transaction is correctly identified as containing FMS funding, that does not necessarily mean that the entirety of the transaction, let alone the contract, are for an international audience. FMS is tightly integrated into the U.S. acquisition system and practitioners have noted that international customers may only be one funder among many in a large bundled buy.

### *Transaction Costs*

The first hypothesis comes directly from the economics literature and the asset specificity theory that if there is a perceived greater and more widespread potential demand for a product, this incentivizes a variety of investments with positive implications for acquisition performance through decreased transactions cost.

> $H_1$ **Lower Specificity:** As the number of export agreements for a project increase (decrease), the likelihood of ceiling breaches and terminations decrease (increase).

In exploring this variable, the study team intends to incorporate, where possible, controls relating to other parts of transaction cost. For example, if there is sufficient data on the use of international supply chains, or, less likely, latitudinal data on offsets, then these variables would be worth including to help distinguish asset specificity from other positive or negative influences on transaction cost.

***Security Partnership***

The next three hypotheses require identification of recipient countries, which the FPDS dataset has not yet achieved. The project team intends to apply machine learning techniques to transaction descriptions. However, that capability is still untested and, even with a hybrid approach including manual identification, may not prove sufficient to the task. The project's back-up plan for recipient identification is to limit the dataset for these variables to those with easily identified projects (e.g., major defense acquisition programs), where the recipients can be more easily determined through other primary and secondary sources including DSCA and SIPRI. In either case, one challenge with this approach is that a given transaction may have a one-to-many relationship with recipients. Once better data is produced, the study team will determine a means of aggregation (e.g., average rating for recipients or looking at the minimum score to identify the weakest link) and will apply this approach consistently across all hypotheses.

The next hypothesis draws more on economic literature than wider foreign policy concerns.

> $H_2$ **Past Deliveries:** As the number of past bilateral deliveries increase (decrease), the likelihood of ceiling breaches and terminations decrease (increases).

This hypothesis posits that more interactions with the U.S. acquisition system will smooth the path for subsequent cooperation, both in terms of building out the bilateral relationship and improving country proficiency with the sometimes arcane U.S. system. This hypothesis intentionally emphasizes the number of transactions rather than the value of those transactions to put small and large countries on similar footing and also because more routine cooperation, even for less valuable items, may show more about the relationship than high profile projects.

The third hypothesis looks at similar questions but through more of a security assistance lens.

> $H_3$ **Alliance Status:** As the recipient's integration into the alliances with the seller increases (decreases), the likelihood of ceiling breaches and terminations decrease (increases).

While formal alliances are clearly demarcated, there are some measurement challenges with this variable. For example, Egypt is a major non-NATO treaty ally but is not part of a formal mutual defense pact with the United States while the Rio Treaty includes a score of Western Hemisphere countries including Cuba, which is not known for its security cooperation with the United States (U.S. Department of State, Treaty Affairs, n.d.). That said, the NATO alliance in particular incorporates both collective defense measures and acquisition related provisions and thus some level of differentiation should be possible, perhaps along separate treaty commitment and defense acquisition arrangement axes.

The last hypothesis has perhaps the strongest theoretical justification in the security assistance literature, but will also be a challenge for measure identification.

> $H_4$ **Institutional Health:** As the health of the recipient's institutions increase (decrease), the likelihood of ceiling breaches and terminations increases (decreases).

This hypothesis has multiple justifications. In political economy terms, more robust institutions reduce the risk of corruption and suggest greater capacity on the recipient's part

and a lower risk of process breakdown. Second, institutional strength may prove directly relevant to the Leahy Laws, that restrict arms transfer to units in purchasing countries with a history of human rights abuses. The most direct justification comes from Andrew Boutton (2018), who argues that "that in uncertain political environments—where regimes are driven by internal power struggles and institutions are underdeveloped—military aid can create a dangerous moral hazard" (p. 7). Recipients who believe that their relationship with the provider ensures their security may engage in coup-proofing behavior that undermines the effectiveness of military institutions and may exacerbate grievances within their country.

## Data and Methods

### *Data Sources and Structure*

#### *Identifying the Datasets*

This paper identifies three primary datasets for studying FMS. The first of these is the Defense Security Cooperation Agency's (DSCA's) *Historical Facts Book*, which provides country-level overviews for arms transfers (2017). This data is available in PDF form, which our team scraped to assemble a dataset tracking country-year level data for FMS agreements and deliveries from 2010 to 2017. The DSCA data does not provide data on individual transactions.

The second dataset is the Stockholm International Peace Research Institute's Arms Transfer database (SIPRI, 2019). SIPRI provides as complete a record as possible of all international arms transfers, including transfers performed via direct commercial sales and transfers from providers other than the United States. SIPRI does not include services. Unlike DSCA, SIPRI provides information on individual transfers, including platform and delivery date. Importantly, due to the variability in pricing between identical platforms, SIPRI does not attempt to provide transaction size in U.S. dollars, but uses a custom Trend Indicator Value (TIV) metric. TIV captures the military significance of the hardware transferred, and is intended for capturing general trends in transfers, not for measuring the discrete dollar size of the transfer. This limits the ability of SIPRI data to be directly integrated with other sources, but it provides extremely valuable directional data on transfers at both the country and platform level.

The final and most substantial dataset is the Federal Procurement Data System's database of all acquisition transactions which use the federal procurement system. FPDS offers extremely granular data on transactions, allowing for detailed breakdowns along types of contract structures, platforms, level of competition, and similar variables. Whether or not a transaction is FMS is recorded in the "foreign funding" field which "indicates that a foreign government, international organization, or foreign military organization bears some of the cost of the acquisition" (USA Spending, 2019). While FPDS provides by far the most granular data on transactions, it suffers two major drawbacks. First, it does not provide explicit information on recipient countries, although some degree of country-attribution may be extracted from plaintext descriptions. Secondly, as shown in ,Figure below, "foreign

funding" is only reliably recorded in recent years, with a majority of data before 2015 unlabeled.

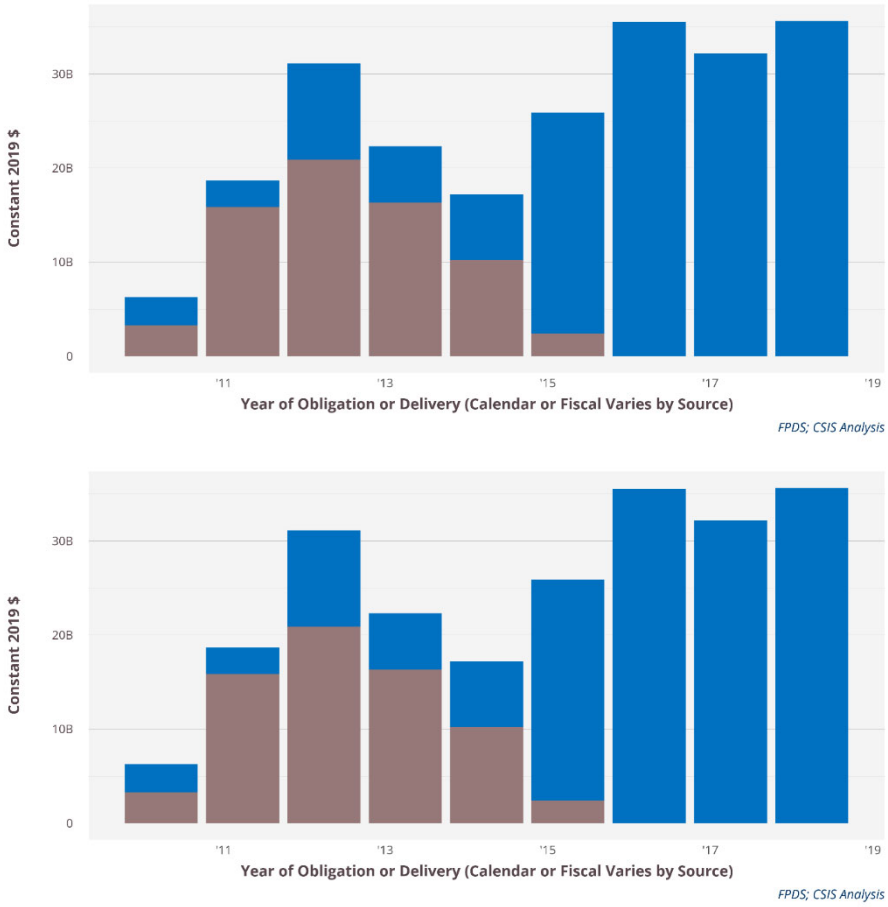


**Figure 2. Limitations of Labeling of Foreign Funding**

***Machine Learning***

Working with the FPDS data for analyzing FMS involved a significant challenge with missing data. While 2016, 2017, and to a lesser extent 2015 were all reliably coded for foreign funding, in previous years coding was sparse or non-existent. In order to extend any analysis prior to 2015, it will be necessary to create some form of classification process, in which unlabeled FPDS transactions can be classified as either FMS or non-FMS.

To classify the unlabeled historical data, we tested three different classification approaches. The first was a simple set of handwritten rules, in which transactions were labeled based on their agency and treasury account codes. The remaining two classifiers were both machine learning approaches, in which a machine learning algorithm was trained on several million labeled transactions to learn patterns to use in classifying new unlabeled observations. The first machine-learning approach uses a random forests algorithm, which creates a large number of decision trees and aggregates their predictions (Breiman, 2001). The second approach uses deep learning, which creates a series of artificial "neurons"' capable of learning complex patterns (LeCun, Bengio, & Hinton, 2015).

All three classification methods were developed using labeled data from FY 2016 and FY 2017, and were then tested on the entirety of the labeled FY 2015 data. The performance from the hand-rules and random forest models are shown below. Precision

captures the rate of true positives from classification (i.e., when the classifier predicts something as FMS, what percentage of the time is it correct?). Recall measures the number of cases captured by the classifier (i.e., what percentage of all FMS transactions did the classifier correctly predict were FMS?). F1-score is the harmonic mean of precision and recall and is a standard overall measure of classifier performance. In all three cases, a score of 1 indicates a perfect classifier.

As shown in Figure 3, the random forest and hand-coded classifiers have similarly strong performance, with manual coding showing slightly greater precision while random forest performs better in recall. Both models correctly label the majority of cases, particularly measured by dollars. So far, deep learning has failed to generalize to the 2015 data, but strong results in the initial 2016/2017 test data indicate potential for improved performance. Existing literature has found that deep learning models outperform random forest models for high-cardinality datasets like FPDS (Guo & Berkhahn, 2016).

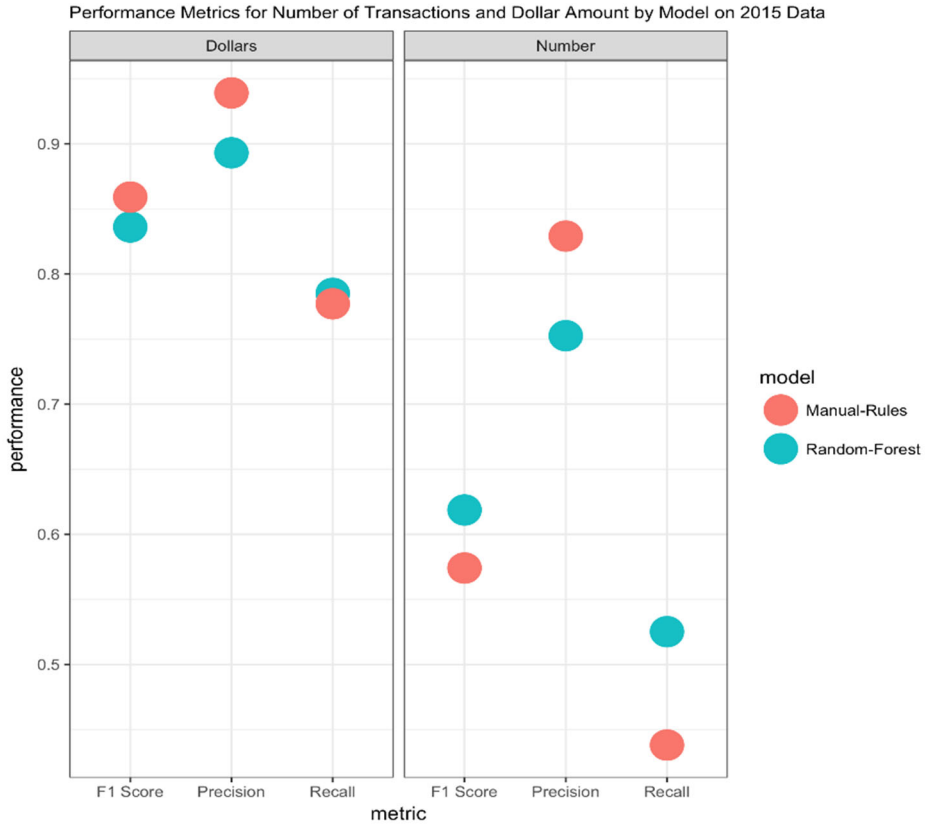


**Figure 3. Classifier Performance**

One additional complication of the FPDS data is the substantial variation in transaction size. Single transactions can range from thousands of dollars to billions of dollars. Because of this variation, it is generally more useful to perform statistical analysis with total dollar figures, rather than transaction counts. However, machine learning algorithms and performance metrics typically operate at the level of observations, not at aggregated values from those observations. In practice, this means that the classifiers discussed here train and measure success based on the number of transactions they correctly classify, not the number of dollars they correctly classify. This volatility in transaction significance makes a strong case for a human-machine hybrid approach, in

which a machine learning algorithm is used in a first pass to classify all transactions, and then a human verifies the few largest transactions by hand. In practical terms, this is only possible for transactions that the classifier identifies as FMS; because the vast majority of all transactions are classified as non-FMS, an inordinate amount of hand-vetting would be needed to cause any meaningful change in performance. For positive predictions, however, limited human-checking of the largest predictions can prevent costly false positives.

All three classification methods offer distinct strengths moving forward. The handwritten rules and random forest models already provide strong performance in classifying both transactions and dollars. The handwritten rules have the advantage of being simple and fast to implement, but they are also inflexible, incapable of using more complicated forms of data, and offer limited room for future improvement. Both the random forest and deep learning methods are more complex and more difficult to implement but offer considerable flexibility and room for future improvement. Both models have the capacity to incorporate plain-text descriptions of contracts and add other variables as desired. Deep learning models offer greater flexibility in incorporating text and are able to capture more complex relationships but are less interpretable than the random forest models and thus far have delivered worse performance on historical data. In their current state, random forests deliver the best overall combination of performance, flexibility for future improvement, and interpretability of results.

Future work on classification strategies will include the incorporation of plain-text descriptions of contracting requirements, which should improve performance. In addition, the current handwritten rules rely entirely on treasury account information, which is not available prior to 2012 in a usable format. Both machine classifiers can generate predictions without using that information, though issues of changing offices and similar new-data problems increase for any classification strategy as it moves back in time. Ideally, either random forest or deep learning models will be able to reliably extend FMS classification significantly past 2012. Together, these classification strategies provide methods for significantly expanding existing FMS datasets and enabling granular historical analysis of FMS transactions.

***Measurement of Independent and Dependent Variables***

While all three datasets have limitations, between them they offer a number of valuable measures for analyzing FMS. DSCA, SIPRI, and classifier-extended FPDS allow for analysis of high-level trends in FMS expenditures over the last several years. These trends are measured via dollar value of total obligations for FPDS, dollar value of FMS delivers for DSCA, and in TIVs for SIPRI.

Both SIPRI and FPDS record the type of arms being transferred, allowing for platform-level breakdowns of trends. However, the two datasets use a different taxonomy of platforms; for instance, SIPRI includes engines as a separate category while FPDS does not. Our team assembled a crosswalk from SIPRI to FPDS by coding the individual weapons platforms in the SIPRI categories which did not match FPDS portfolios. This makes it possible to breakdown SIPRI data into FPDS portfolios, allowing comparison between SIPRI and FPDS at the platform level. Additional work will be required to enable FPDS-to-SIPRI translation.

To analyze the characteristics of FMS transactions, we use a simplified version of FPDS's "Type of Contract" field, which indicates the use of fixed fees, incentives, and so forth. This allows for comparison between FMS transactions and other non-FMS DoD transactions in terms of which pricing mechanisms they use. Future work on contract

performance may draw on FPDS's measure of terminations, cost-ceiling breaches, and extent of competition.

## Preliminary Results

### *Comparing Sources*

All three datasets exhibit some level of agreement at the year level on general trends in FMS. However, there is a large discrepancy between FPDS and DSCA totals, with FPDS figures generally substantially exceeding DSCA figures. This is especially surprising given the lower quality of labeling in the FPDS dataset. This may be in part due to partially-FMS transactions being labeled as wholly-FMS by FPDS, though hand-verification of the largest FMS transactions in FPDS did not find any mixed transactions. There is also some issue of timing: FPDS, for instance, contains several large FMS transactions in 2012 due to obligations to produce a number of aircraft, while DSCA deliveries may smooth those obligations out as the aircraft are delivered over several years and tend to occur later in time.

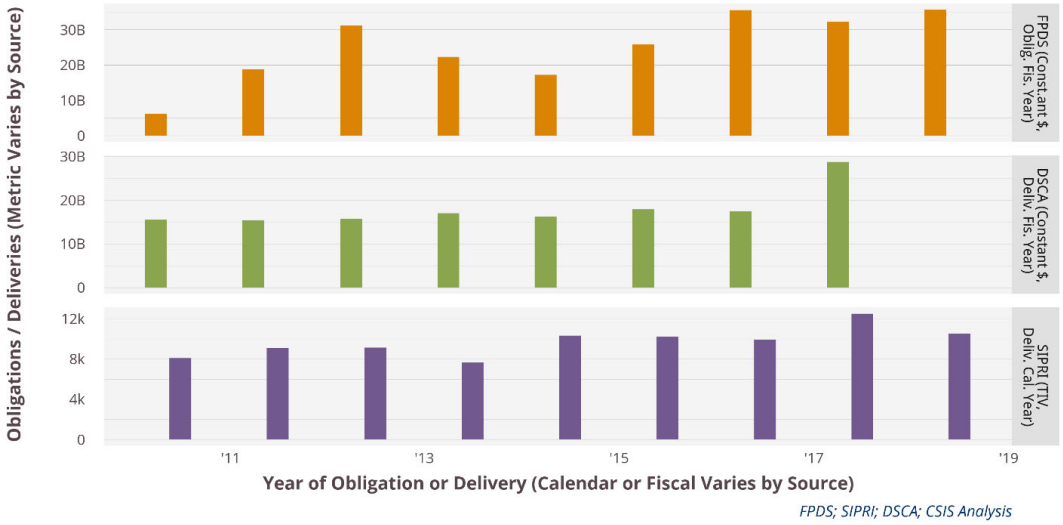


**Figure 4. Annual Funding by Source**

While DSCA and FPDS show some agreement on trends, if not precise amounts, SIPRI appears to deviate from both DSCA and FPDS in year-to-year trends. Some of this is due to the nature of how SIPRI calculates TIVs, however. As shown in Figure 5 breaking out FPDS and SIPRI by platform shows a much greater degree of agreement. Both FPDS and SIPRI show that aircraft dominate U.S. arms exports. They show similarly low and relatively steady rates for ships and submarines and land vehicles. Both FPDS and SIPRI show a sharp and steady increase in Ordinance and Missile exports, though they disagree on the trends for sensors. On the whole, the platform-level analysis reveals a high level of agreement between SIPRI and FPDS, with the disagreements on the aggregate level appearing to be primarily a result of different calculations of aircraft value, possibly due to TIV calculations for the Joint Strike Fighter.

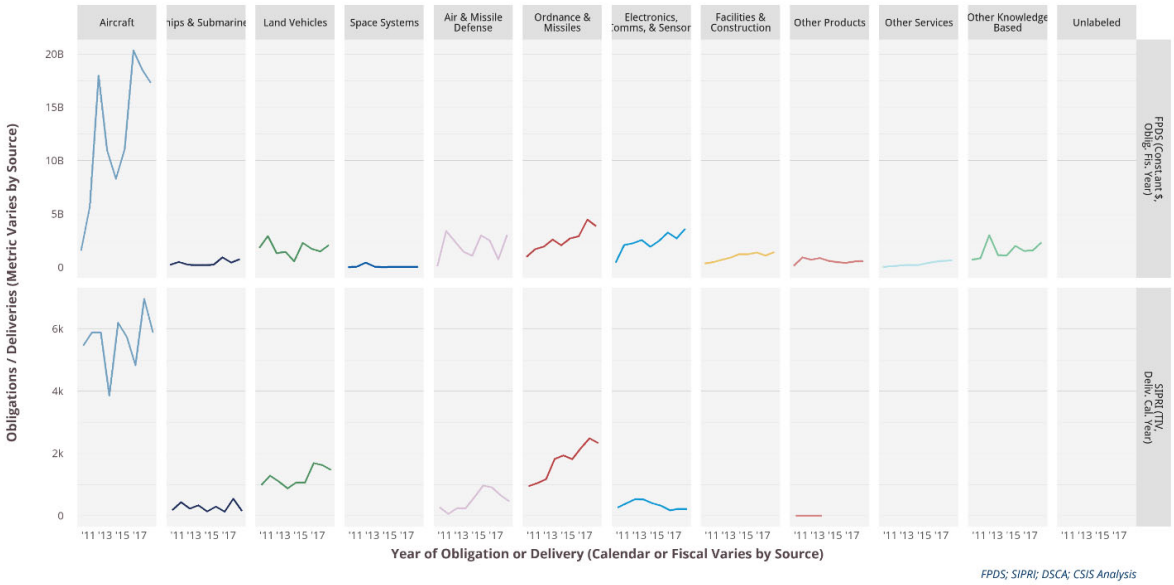


**Figure 5. Comparing FPDS and SIPRI**

### *Contracting Approaches*

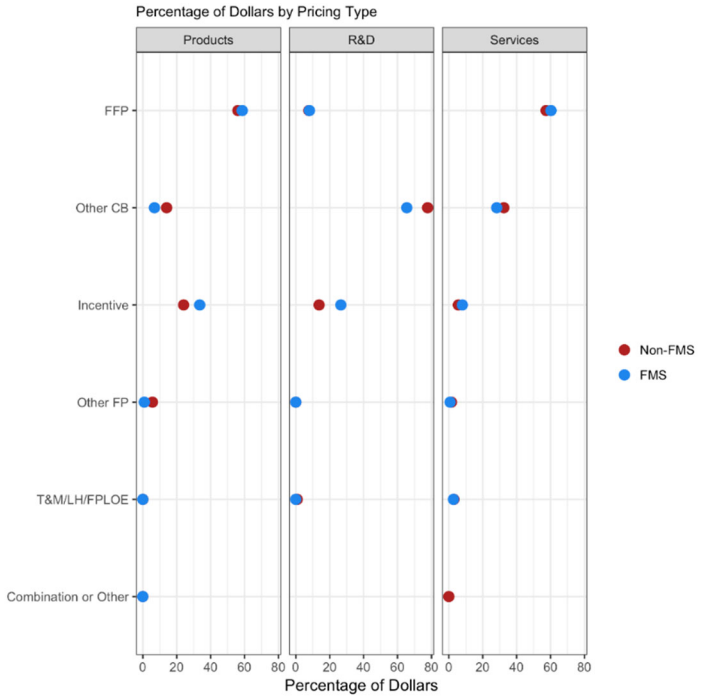


Figure 6 shows the breakdown of contract pricing types for FMS and non-FMS DoD transactions. FMS and non-FMS pricing structures are similar in many ways, especially for service provision. For both products and R&D however, there is meaningful divergence in contract structure in keeping with expectations from the theoretical literature. FMS transactions tend to use incentive-based contracts, specifically fixed-price incentive fee, more frequently than non-FMS transactions. That approach was favored, where appropriate, by the Better Buying Power initiatives and would be in keeping with the use of higher-incentive contracts in the presence of reduced monitoring capacity and higher transaction costs as may be the case in

international transfers. Interestingly, the higher use of incentives by FMS contracts does not result in a drop of firm-fixed-price contracts. Instead, FMS transactions tend to use other cost-based mechanisms less often than non-FMS transactions, which may suggest differences in monitoring capacity or degree of trust for domestic sales as opposed to FMS.

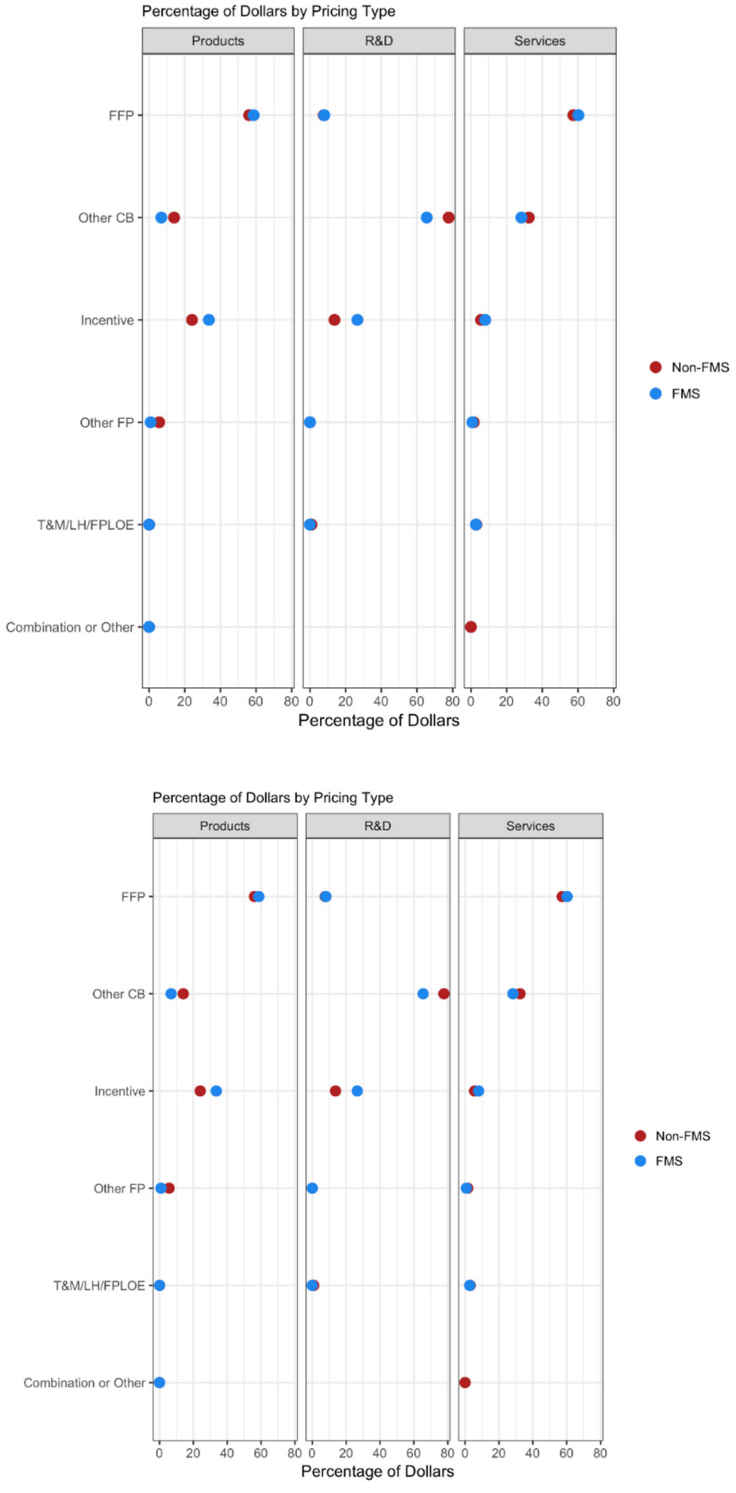


**Figure 6. Contract Pricing for FMS Versus Non-FMS Contracts**

## Discussion and Next Steps

The limitations on tracking FMS spending in FMS significantly impede not just the research questions raised in this paper, but a range of other pertinent questions regarding this important and controversial subset of defense contracting. For security sector assistance in particular, assessment, monitoring, and evaluation are watchwords. More rigorous data enables anyone seeking to understand the benefits and risks of present FMS. The biggest surprise thus far in the results is that both in the years that are best labeled, as well as in those that are likely missing some FMS contracts, FPDS obligation levels exceed the deliveries reported by DSCA. The study team will look closely at this issue, examining issues of the timing of obligations versus deliveries, as well as the bundling of FMS and domestically-funded transactions.

This project still has important steps ahead, particularly in the further integration of FPDS, SIPRI, and DSCA data across country and project lines where possible. Taking those steps will better enable the analysis of hypotheses, improves the study team's ability to validate FPDS obligation levels, and enables future researchers and practitioners seeking to better understand the interconnected and high stakes issues surrounding FMS.

## References


Acquisitions for Foreign Military Sales, 48 U.S.C. §225.7306.

Adler, T. R., Sherer, R. F., Barton, S. L., & Katerberg, R. (1998). An empirical test of transaction cost theory: Validating contract typology. *Journal of Applied Management Studies, 7*(2), 185–200.

Boutton, A. (2018). The dangers of U.S. military assistance to weak states. *Policy Roundtable: The Pros and Cons of Security Assistance, Texas National Security Review,* 7–18.

Breiman, L. (2001). Random forests. *Machine Learning*, *45,* 5–32.

Brown, T. L., & Potoski, M. (2003). Transaction costs and institutional explanations for government service production decisions. *Journal of Public Administration Research and Theory, 13*(4), 441–468. Retrieved from https://www.jstor.org/stable/3525657

Censer, M. (2018, June 29). Pentagon issues final rule on FMS offsets. Retrieved from https://insidedefense.com/insider/pentagon-issues-final-rule-fms-offsets

Dalton, M. (2018). The risks and tradeoffs of security cooperation. *Policy Roundtable: The Pros and Cons of Security Assistance, Texas National Security Review,* 35–41.

Dalton, M., Shah, H., Green, S. N., & Hughes, R. (2018). *Oversight and accountability in U.S. security sector assistance: Seeking return on investment*. Retrieved from Center for Strategic and International Studies website: https://csis-prod.s3.amazonaws.com/s3fs-public/publication/180207_Dalton_OversightAccountability_Web.pdf

Daniels, S. P. (2018). *Show me the money: Assessing the fiscal reality of the National Defense Strategy's ambitions.* Retrieved from Center for Strategic and International Studies website: http://defense360.csis.org/wp-content/uploads/2018/09/Daniels_Endgame_D360.pdf

Defense Security Cooperation Agency. (2012). *Security assistance management manual: Chapter 5–FMS case development.* Retrieved from https://www.samm.dsca.mil/chapter/chapter-5

Defense Security Cooperation Agency. (2017). *Historical facts book.* Retrieved from https://www.dsca.mil/sites/default/files/historical_facts_book_-_30_september_2017.pdf

Gilman, D. (2014). *Foreign military sales*. Retrieved from Defense Security Cooperation Agency website: https://www.dsca.mil/sites/default/files/final-fms-dcs_30_sep.pdf

De Vore, M. R. (2011). The arms collaboration dilemma: Between principal-agent dynamics and collective action problems. *Security Studies, 20*. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/09636412.2011.625763

Fergusson, I. F., & Kerr, P. K. (2017). *The U.S. export control system and the export control reform initiative.* Retrieved from Congressional Research Service website: https://fas.org/sgp/crs/natsec/R41916.pdf

Guo, C., & Berkhahn, F. (2016). *Entity embeddings of categorical variables* (arXiv:1604.06737v1). Retrieved from https://arxiv.org/abs/1604.06737

Harrison, T., & Daniels, S. P. (2018). *Analysis of the FY 2019 defense budget.* Retrieved from Center for Strategic and International Studies website: http://defense360.csis.org/wp-content/uploads/2018/09/180917_Harrison_DefenseBudget2019_WEB_FINAL_FINAL.pdf

Interagency Task Force. (2018). *Assessing and strengthening the manufacturing and defense industrial base and supply chain resiliency of the United States: Report to President Donald J. Trump by the Interagency Task Force in fulfillment of Executive Order 13806*. Retrieved from https://media.defense.gov/2018/Oct/05/2002048904/-1/-1/1/ASSESSING-AND-STRENGTHENING-THE-MANUFACTURING-AND%20DEFENSE-INDUSTRIAL-BASE-AND-SUPPLY-CHAIN-RESILIENCY.PDF

Kirkpatrick, D. (2004). Trends in the cost of weapons systems and the consequences. *Defence and Peace Economics*, *15*(3), 259–273. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/1024269032000123203

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521,* 436–444.

Little, T. D. (2017, September 5). End-use monitoring is the key to success in foreign military sales. *Army Sustainment Magazine*. Retrieved from https://www.army.mil/article/192447/end_use_monitoring_is_the_key_to_success_in_foreign_military_sales

McCormick, R., Hunter, A. P., & Sanders, G. (2017). *Measuring the impact of sequestration and the drawdown on the defense industrial base.* Retrieved from the Center for Strategic and International Studies website: https://csis-prod.s3.amazonaws.com/s3fs-public/publication/180111_McCormick_ImpactOfSequestration_Web.pdf?A10C65W9Qkx07VaJqYcJguCH.7EL3O7W

Peltzman, S. (1977). The gains and losses from industrial concentration. *Journal of Law and Economics,* 20. Retrieved from https://www.nber.org/papers/w163

Petty, F. S. (1999). *Defense offsets: A strategic military perspective*. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a363929.pdf

Sanders, G., & Cohen, S. (2017). *Designing and managing successful international joint development programs.* Retrieved from the Center for Strategic and International Studies website: https://csis-prod.s3.amazonaws.com/s3fs-

public/publication/170125_Sanders_IntlJointDevelPrograms_Web.pdf?6CJdjahdu6eaSqqv0vZnfioPn.nrHcgr

Sanders, G., & Huitink, Z. (2019). *Evaluating consolidation and the threat of monopolies within industrial sectors.* Retrieved from the Center for Strategic and International Studies website: https://csis-prod.s3.amazonaws.com/s3fs-public/publication/190208_Sanders_Monopolies_WEB_v2.pdf

Schmitz, P. W. (2001). The hold-up problem and incomplete contracts: A survey of recent topics in contract theory. *Bulletin of Economic Research, 53*(1).

Stockholm International Peace Research Institute (SIPRI). (2019). *Sources and methods.* Retrieved from https://www.sipri.org/databases/armstransfers/sources-and-methods#Contents

Trump, D. J. (2018, April 19). *National security presidential memorandum regarding U.S. conventional arms transfer policy* [Memorandum]. Washington, DC: White House. Retrieved from https://www.whitehouse.gov/presidential-actions/national-security-presidential-memorandum-regarding-u-s-conventional-arms-transfer-policy/

USA Spending. (2019). *Data dictionary.* Retrieved from http://files.usaspending.gov/docs/DATA+Transparency+Crosswalk.xlsx

U.S. Department of State, Bureau of Public Affairs: Office of Press Relations. (2018, April 19). *Briefing on updated conventional arms transfer policy and unmanned aerial systems (UAS) exportpolicy* [Press release]. Retrieved from https://www.state.gov/r/pa/prs/ps/2018/04/280613.htm

U.S. Department of State, Treaty Affairs. (n.d.). U.S. collective defense arrangements. Retrieved from https://www.state.gov/s/l/treaty/collectivedefense/

Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology, 87*(3), 548–577. Retrieved from http://www.jstor.org/stable/2778934


ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CA 93943

www.acquisitionresearch.net