Costello College of Business
GREG AND CAMILLE BARONI CENTER FOR GOVERNMENT CONTRACTING
George Mason University

White Paper Series
NO. 22 JUNE 10, 2025

The Potential Role of AI in Legislative Research and Drafting

Results of a Pilot Program

Richard Beutel, Art Nicewick

SYNOPSIS

The integration of artificial intelligence (AI) into the legislative process offers transformative potential for congressional and industry staff tasked with researching, drafting, and preparing legislation. As the complexity of modern governance increases on issues like climate change, technology regulation, and healthcare reform, so too does the demand for precision, efficiency, and clarity in legislative work.

AI tools, such as those leveraging natural language processing (NLP), machine learning, and data analytics, can assist analysts in overcoming longstanding challenges in legislative drafting, evaluating legal precedents, amending existing frameworks, and producing explanatory statements and committee notes. However, their adoption also raises questions about accuracy, accountability, and the preservation of human judgment in lawmaking.

To explore these issues, the Baroni Center undertook a pilot effort to train two Large Language Models (LLM's), Chat GPT and GROK, and applied them to one of the most ambitious legislative reform packages in recent years—the FoRGED Act from Senator Roger Wicker. The FoRGED Act[1] (Fostering Reform and Government Efficiency in Defense Act) is an enormously complex reform of the entire defense acquisition procurement system. It will, when enacted, transform the fashion in which billions of dollars of defense procurement is organized and conducted.

This report describes the results of our experiment. The Appendix provides the link to the analysis[2] of the FoRGED Act that was created using the two LLM's.

Having trained the LLM's, the results were curated by an expert panel (including members of the Senate Armed Services Committee) in order to assess the results for accuracy, timeliness and thoroughness.

The results were intriguing and hopeful. But concerns did emerge.

> Those curating the AI analysis noted that the AI was reliant on the existing web knowledge base and existing industry and government analysis. That knowledge base often reflected implicit and explicit biases and assumptions. It also could not analyze sources and references that did not yet exist.

> One curator noted that we are asking the AI to answer questions that it has insufficient information to answer.

> Another commentator noted that the AI provided what it has and makes it sound like it has an authoritative answer—when it doesn't.

> Rather than thinking about questions we want to have answered, what does this provision do? Is it a good provision? AI trainers should think about what questions the AI is likely to have sufficient information to appropriately answer.

We observed that one of the most important things that a researcher must do is to caveat his/her findings by acknowledging what exactly he/she doesn't know. The absence of such caveats from this model may be more problematic than the absence of existing sourcing. Can AI be trained to discuss what it doesn't know?

As part of this pilot effort, we have provided our hosted link (see Appendix) of the AI-based legislative analysis to the general public. We make no assurances as to how this AI model is or should be utilized. Anyone using the model does so on their own responsibility.

Our analysis could support the public policy dialogue and inform the outcome of this important legislative package.

1. Fostering Reform and Government Efficiency in Defense Act or "FoRGED Act." S. 5618,118th Congress (2024). https://www.congress.gov/bill/118th-congress/senate-bill/5618/text/is.

2. AI Pilot analysis host site, https://forged-act.com/.

USING AI FOR LEGISLATIVE ANALYSIS: IDENTIFY THE CONCERNS TO BE ADDRESSED

In structuring this AI pilot, we began by working with a data scientist from Defense Acquisition University to identify specific concerns that any AI analysis of legislative proposals would need to address.

Legislative drafting is a meticulous process requiring clarity, consistency, and foresight to avoid ambiguity or unintended consequences. Analysts creating complex legislation face a daunting task—they must begin by crafting language that aligns with constitutional principles, statutory intent, and practical enforceability, all while navigating tight deadlines and political pressures.

We wanted to determine if AI could streamline this process by analyzing the often-significant repositories of existing laws, regulations, and judicial interpretations to suggest precise phrasing. For instance, could LLM platforms identify ambiguous terms or conflicting provisions within a draft, flagging them for revision before they move through regular order?

Clearly, AI platforms face major challenges. How do AI systems address the nuanced, context-dependent nature of legislative intent and avoid being misaligned with political or ethical goals? Does human oversight remain essential to ensure that AI-generated drafts reflect the values and priorities of elected representatives?

EVALUATING LEGAL PRECEDENTS WITH AI

Another area we wanted to explore was whether AI could identify and explain legal precedents and the broader policy context. Understanding legal precedents is critical to understanding legislation, as new laws must fit within the existing judicial landscape.

Traditionally, manual research is required to identify relevant case law, a time-intensive process prone to oversight. Can AI revolutionize this by rapidly analyzing thousands of court decisions, statutes, and regulatory rulings to pinpoint applicable precedents?

We also wanted to see if GROK and ChatGPT could be used to detect patterns in judicial reasoning, so as to highlight how courts have interpreted similar provisions in the past and predicting potential legal challenges to proposed legislation.

We know that the reliance on AI for precedent evaluation introduces risks. Algorithms may overemphasize quantitative patterns (e.g., frequency of citations) over qualitative factors (e.g., judicial philosophy), thereby skewing the analysis. Analysts must, therefore, verify AI outputs against primary sources to ensure that the technology complements rather than supplants human expertise.

WHY DID WE USE THE FORGED ACT?

To structure this pilot, we needed to find a suitable legislative framework in which to apply our LLM's.

The FoRGED Act (Fostering Reform and Government Efficiency in Defense), formally introduced as S. 5618 by Senator Roger Wicker (R-MS) on December 18, 2024, is a comprehensive legislative proposal aimed at reforming the U.S. Department of Defense (DoD).

As the ranking member of the Senate Armed Services Committee, Senator Wicker developed this legislation to address inefficiencies, enhance national security, and modernize Pentagon operations.[3]

The stated objectives of the FoRGED Act include:

- **Maximizing Taxpayer Dollars**: The legislation seeks to optimize defense spending by improving procurement processes, ensuring funds are used effectively to acquire cutting-edge weapons systems.
- **Accelerating Technology Adoption:** It aims to streamline the integration of new technologies into DoD operations, reducing bureaucratic delays that hinder innovation.
- **Boosting Competition**: The Act encourages competition within the defense industry to drive quality and cost-efficiency, potentially by revising contracting practices.
- **Enhancing National Security**: By reforming how the DoD operates, the legislation intends to strengthen deterrence against adversaries and restore public confidence in military spending.

The FoRGED Act reflects Wicker's ongoing emphasis on national security and efficiency, building on his prior legislative efforts like data privacy reforms. As such, it provided us with an excellent use case in which to test our AI LLM's.

This is because, *inter alia*, the FoRGED Act requires precise language to reform DoD processes without disrupting ongoing operations.

3. The FoRGED Act complements Senator Wicker's earlier report, "21st Century Peace Through Strength: A Generational Investment in the U.S. Military," which called for significant Department of Defense investment to counter threats, particularly from China. https://www.wicker.senate.gov/2024/5/senator-wicker-unveils-major-defense-investment-plan.

HOW DID WE TRAIN THE TWO LLM'S TO TACKLE THE FORGED ACT?

To apply Chat GPT and GROK to the FoRGED Act, we trained these platforms to examine the text of the legislation with the following queries.

> Did the legislation utilize consistent terms? AI tools can scan drafts for ambiguous terms (e.g., "efficiency" or "competition") and suggest definitions or uniform usage, aligning with legal standards like those in the National Defense Authorization Act (NDAA).
>
> Did the FoRGED Act amend titles such as 10 U.S.C. (Armed Forces), and if so, what was the background of these repealed provisions and what was the likely consequence or second order effects of these repeals? Could AI automatically identify and link relevant sections, reducing manual errors and ensuring seamless integration?

We also wanted to know if the LLM's could analyze past defense reforms and observe whether they succeeded or failed. Specifically:

- **Analyze Historical Legislation**: Compile a database of germane prior defense bills (e.g., past NDAA's) and judicial rulings (e.g., on defense contracting disputes), and highlight precedents for restructuring or technology adoption.
- **Predict Legal Challenges**: Identify potential vulnerabilities to simulate how courts might interpret FoRGED Act provisions based on cases like *United States v. Lockheed Martin* (contract disputes).
- **Benchmarking**: Compare the FoRGED Act to International Traffic in Arms Regulation (ITAR) or North Atlantic Treaty Organization (NATO) procurement standards, offering insights into international defense legal frameworks.

We were also interested in whether the FoRGED Act likely modifies existing defense laws, thus requiring careful integration. For example:

- Could the analysis of the FoRGED Act map internal dependencies?
- Could it create a "dependency tree" of affected statutes (competition requirements), showing how changes ripple through the U.S. Code?

Was it capable of analyzing current inefficiencies?

- Like delays in the DoD's Joint All-Domain Command and Control (JADC2) program and suggest targeted revisions to acquisition rules?
- Would this detect unaddressed areas, such as cybersecurity for new tech, and recommend additional provisions to align with policies like the Cybersecurity Maturity Model Certification (CMMC)?

THE STRUCTURE OF OUR TRAINING APPROACH

We designed the pilot by training the LLM's to undertake the following:

- **Synthesize Research**: Summarize DoD reports, GAO audits (e.g., on weapons system delays), and Wicker's "Peace Through Strength" report into concise rationales for reform.[4]
- **Draft Notes**: Show results using a spreadsheet format that was adaptable to such uses as preparing initial committee notes tailored to audiences like the Senate Armed Services Committee.
- **Enhance Accessibility**: Produce plain-language summaries for constituents.

In technical terms, the FoRGED Act analysis for Title One was developed through the use of two large language models and a series of prompts crafted via excel to ask targeted questions. The Excel sheet maintained a different recommendation of repeal action in the first column of the spread sheet, then the subsequent columns merged that code into a complex question.

Using this simple process, we were able to construct a series of complex prompts to request information about the FoRGED Act provisions.

Using this simple approach, we created an Excel sheet of the 283 repeals listed in the FoRGED Act, then constructed 10 questions for each repeal, for a total of 2,830 Prompts. Next, we selected the 30 recommendations, constructed 4 questions for each of these, creating an additional 120 prompts, for a total of 2,950 prompts.

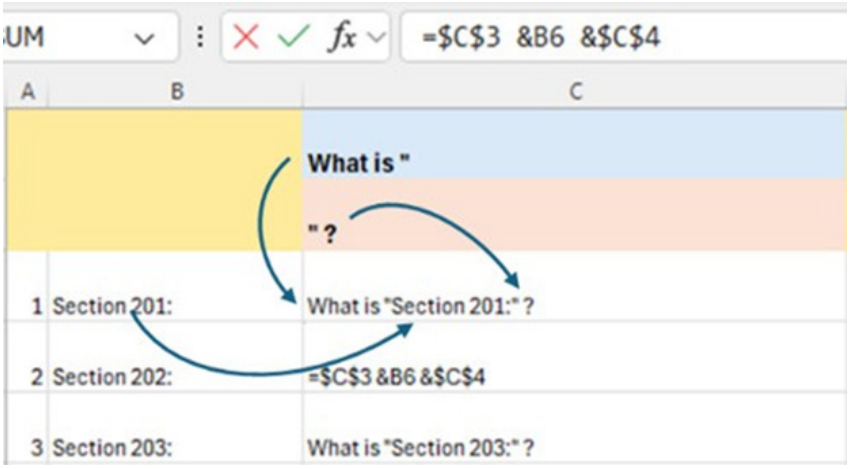
We tested the execution of this batch of prompts on multiple platforms and discovered that the quality of the response using the "Deep Research" option on the LLM was greatly enhanced through the use of advanced reasoning. However, most platforms run into performance limitations.

Through a series of tests, we observed that the GROK3 model provided high quality results, at a greatly increased performance level, and in some cases 10X the speed of the competitors. This may be related to the huge clustering used in the XAI implementation. (See Figure 1. Process Diagram)

Despite these advantages, AI-generated explanatory materials may lack the persuasive finesse or rhetorical flourish that human drafters bring to bear, particularly in politically charged contexts. Moreover, errors in AI summarization could misrepresent intent, necessitating rigorous proofreading by staff.

4. 21st Century Peace Through Strength.

Figure 1. Process Diagram



Defined the desired effects of *Section 201: Transition of Program Executive Officer Role to Portfolio Acquisition Executive*, *list out the potential positive effects of this provision. Reference the Forged Act located at https://www.congress.gov/bill/118th-congress/senate-bill/5618/text/is?-format=txt. Keep response under 500 words*.

Bold: "Before Text"
Bold Italic: the Section
Italic: the "After text"

FEEDBACK FROM THE PANEL CURATING THE ANALYSIS

To assess the quality or utility of our analysis of the FoRGED Act, we organized a panel of defense acquisition policy professionals to assess, curate, and critique the results of the analysis.

Their feedback was informative. Those curating the AI analysis noted that the AI was reliant on the existing web knowledge base and existing industry and government analysis. That knowledge base often reflected implicit and explicit biases and assumptions. It could not analyze that which did not yet exist.

- One curator noted that we are asking the AI to answer questions that it has insufficient information to answer.
- Another commentator noted that the AI made it sound like it has an authoritative answer—when in fact it does not.
- Rather than think about questions to be answered (What does this provision do? Is it a good provision?), AI trainers should think about what questions the AI is likely to have sufficient information to appropriately answer.

We observed that one of the most important things that a researcher must do is to caveat his/her findings by acknowledging what he/she doesn't know. The absence of such caveats from this model may be more problematic than the absence of sourcing. Can AI be trained to discuss what it does not know?

Another curator liked the video, and particularly the analogies. They noted that the Title 3 [Rapid Acquisition and Commercial Contracting] analysis was superficial in terms of its legislative goals and objectives. But they indicated that the videos were very general and suited for the general public, not subject matter experts.

The analysis of Title 4, Promotion of Competition in the Defense Industrial Base, was problematic. Title 4 did not put more scrutiny on mergers and acquisitions as the AI analysis suggested. It is primarily a Qual/test focus with a focus upon setting aside funds to go after non-responsive sole source suppliers.

Other comments concluded that the AI generated discussion on "sunsetting" of these provisions useful. Note that this idea—the utility of sunsetting provisions—came from the AI itself and was part of our explicit training regime.

Curators concluded that taking a core idea for each of the parts of the FoRGED Act and describing what the Act will do is a useful feature of the anal-

ysis. Curators specifically liked the AI discussion regarding exemptions to DoD unique requirements and separating them from non-traditional contractors. Other comments noted with approval the discussion to put teeth into the commercial item preference by starting with commercial items in DoD procurements and only going "defense unique" with DoD approval based on the market research.

Curators also criticized the last discussion in the AI analysis on the FoRGED Act impacts on the budget process seemed to miss the point. Creators of the FoRGED Act stated to us that DoD needs flexibility within budget lines to move money at a speed that matches the pace of change in threats and technologies. Senate staffers noted that DoD cannot wait around for three years as the world changes, leaving DoD constantly behind the eight ball.

Other comments included:

- "I don't think it is 'complete' and that to me is a major detractor. AI should be able to help staff/legislators understand the scope of a particular issue, gather all the data, and present it for consideration by the author of a piece of legislation before voting on it. Commercial items are a perfect example . . . [the] policy [on commercial items] has moved back and forth and folks need to understand whether, if adopted, the language in the FoRGED Act makes a definitive statutory change. The analysis here doesn't."
- "In a recent study I served on it would have saved a lot of time, and thus dollars, had we been able to do in minutes the background research on the topic."
- "I did not think that using press releases and other public documents to form an analysis was the best possible source, as these are almost always one-sided in Washington when it comes to legislative proposals and any AI use risks capturing that perspective rather than providing an analysis. Perhaps instead of using such "propaganda" to form a general analysis, perhaps you can add a running list of hyperlink materials and sites that AI compiles from proponents and opponents to capture those points of view."
- The "deep dive" link into each section provided a much better summary and explanation of what was being proposed and what the impact would be. I thought this came close to the citations and depth of impact analysis that stakeholders would need, but some sections continue to lack the level of detail.
- "I thought the podcast audio was a well-done summary that could be used to educate the general public (on a legislator's website, for example). The only problem with this presentation format is that none of them were 'short' and if someone wanted to understand the entire proposal, it could be a long audio session."

The Baroni Center, sincerely thanks the expert panel for their detailed curation of this AI pilot effort.

CONCLUSION

AI's ability to process vast datasets (e.g., DoD budgets, contractor performance records) accelerates research, while its predictive tools help anticipate implementation challenges. For instance, our LLM's could model how the Act's reforms impact deterrence against China using unclassified threat assessments. However, staff must oversee AI outputs to ensure alignment with Senator Wicker's intent and political realities—AI might excel at technical drafting but miss nuanced stakeholder dynamics. Ethical use, such as avoiding biased data in procurement analysis, is also critical.

Despite these advantages, AI-generated explanatory materials may lack the persuasive finesse or rhetorical flourish that human drafters bring to bear, particularly in politically charged contexts. Moreover, errors in AI summarization could misrepresent intent, necessitating rigorous proofreading by staff.

Adapting legal frameworks with AI is not foolproof.

- The technology may lack the political acumen to weigh trade-offs or anticipate stakeholder reactions, areas where human judgment remains irreplaceable.
- Additionally, over-reliance on AI could lead to "path dependency," where staff favor incremental tweaks over bold reforms, simply because the system excels at refining what already exists.

RESEARCH CURATORS

Peter Levine: Mr. Levine is one of the foremost experts on the history and complexities of the National Defense Authorization Act, having served as General Counsel to the Senate Armed Services Committee for many years, as well as in senior leadership positions in the DoD.

Trey Hodgkins: Mr. Hodgkins was the former Senior Vice President, Public Sector at the Information Technology Alliance, where he managed a team of professionals responsible for all advocacy, public policy and business development activities for companies facing the Public Sector market. He is currently in private practice.

David Drabkin: Mr. Drabkin is a nationally recognized expert in federal acquisition policy served in senior leadership roles at GSA. He also served as Chairman of the seminal Section 809 Panel. He is currently in private practice.

Certain Professional Staff Members of the Senate Armed Services Committee

APPENDIX

USING LARGE LANGUAGE MODELS TO ANALYZE THE TEXT OF THE FORGED ACT

Link to Forged-Act.com and the AI analysis: https://forged-act.com

How to Use forged-act.com

Forged-Act.com was created after it became clear that a static PDF could not satisfy the diverse needs of readers following the Forged Act. The original AI-assisted study examined all thirty-one sections of the bill and a separate study reviewed the 283 repealed provisions.

For each of the thirty-one sections the paper explored ten analytical dimensions:

1. Key points
2. Legislative history
3. Anticipated benefits
4. Potential unintended consequences
5. Mitigation strategies
6. Affected stakeholders
7. Likely opponents
8. Additional resources required for implementation
9. Metrics for success, and
10. Alternative approaches

Peer reviewers soon identified several gaps. Seasoned analysts wanted deeper, section-specific evidence, while newcomers needed an accessible on-ramp to understand the Act as a whole.

To bridge both ends of that spectrum, the content was migrated to a full website. Each section now includes a "Deep Analysis" page for researchers and an accompanying audio podcast that delivers a easy to consume overview. A video version of each podcast—currently in the works—will offer a visual counterpart for learners who prefer that format.

Because AI-generated material can contain errors (and hallucinations), the site incorporates several safeguards. The web page opens with a disclaimer, a feedback form invites readers to report inaccuracies, and "Judgement" documents placed in the additional materials section using multiple AI LLM in direct comparison to expose discrepancies. These mechanisms, taken together, aim to improve transparency and steadily refine the accuracy of the analysis.

Although the Forged Act is still making its way through Congress and may change, the site is designed to evolve with it. Our goal is to equip DoD and other federal stakeholders with a living, rigorously reviewed resource that clarifies the Act's implications and supports informed modernization of defense acquisition.

THE FORGED ACT

S. 5618
IN THE SENATE OF THE UNITED STATES
December 19 (legislative day, December 16), 2024

Mr. Wicker introduced the following bill; which was read twice and referred to the Committee on Armed Services

A Bill

To promote defense innovation, and for other purposes.

Title I. Short Title

Section 101: Repeals of Existing Law to Streamline the Defense Acquisition Process
Section 102: Modifications to Current Defense Acquisition Requirements
Section 103: Automatic Sunset for Future Statutory Reporting Requirements

Title II. Defense Acquisition Roles, Responsibilities, and Organizations

Section 201: Transition of Program Executive Officer Role to Portfolio Acquisition Executive
Section 202: Amendments to the Joint Requirements Oversight Council
Section 203: Matters Relating to the Director of Cost Assessment and Program Evaluation
Section 204: Establishment of Joint Requirements and Programming

Board
Section 205: Capstone Requirements

Title III. Rapid Acquisition and Commercial Contracting

Section 301: Milestone A
Section 302: Modification to Acquisition Strategy
Section 303: Exemptions for Nontraditional Defense Contractors
Section 304: Modifications to Treatment of Certain Products and Services as Commercial Products and Commercial Services
Section 305: Modification to Nontraditional Defense Contractor Definitions
Section 306: Alternative Capability Based Pricing
Section 307: Modifications to Certain Procurement Thresholds
Section 308: Modifications to Commercial Solutions Openings
Section 309: Modifications to Other Transactions
Section 310: Modifications to Commercial Product and Commercial Service Determinations by Department of Defense
Section 311: Commercially Acceptable Transaction and Payment Methods
Section 312: Transparency and Accountability of Contract Awards
Section 313: Flowdown Requirements
Section 314: Modifications to Relationship of Other Provisions of Law to Procurement of Commercial Products and Commercial Ser**vices**
Section 315: Nontraditional Defense Contractor Commercial Solutions Opening
Section 316: Program Management Office Competition
Section 317: Middle Tier of Acquisition for Rapid Prototyping and Rapid Fielding
Section 318: Revision and Codification of Software Acquisition Pathways
Section 319: Modifications to Steps to Identify and Address Potential Unfair Competitive Advantage of Technical Advisors to Acquisition Officials
Section 320: Modifications to Procurement for Experimental Purposes
Section 321: Consumption-Based Solutions

Title IV. Promotion of Competition in the Defense Industrial Base

Section 401: Review of Structure of the Budget and Appropriations for Funding of Defense Acquisition Programs
Section 402: Administration of the Industrial Expansion Program

Title V. Defense Budgeting Processes

Section 501: Review Process
Section 502: Modifications to the Defense Modernization Account
Section 503: Amendments and Repeals to Budgetary Requirements for Defense Acquisition

Baroni White Paper Series

The purpose of the Baroni White Paper Series is to promote research and discussion on topics of interest and importance to the government contracting community. Comments from the community are welcome and may be sent to govcon@gmu.edu.

NO. 1. *Unintended Consequences of Small usiness Contracting,* Craig R. Reed, Ph.D. November 25, 2019

NO. 2. *Pricing Intellectual Property in Defense Competitions: Toward Theoretical and Practical Advice for Government Officials and Government Contractors,* James Hasik, Ph.D. November 25, 2016

NO. 3. *The Cost of Saving Money: The Negative Impact of Roller Coaster DoD Funding,* Jennifer Taylor. November 25, 2016

NO. 4. *The Value of Intellectual Property in Governm*ent Procurement Auctions, James Hasik, Ph.D. July 14, 2020

NO. 5. *The DoD Budget Process: The Next Frontier of Acquisition Reform,* Eric Lofgren. July 29, 2020

NO. 6. *Building Resilience: Mobilizing the Defense Industrial Base in an Era of Great-Power Competition.* Jerry McGinn, Ph.D. September 28, 2020

NO. 7. *What Future for Remote Work in Federal Contracting?* James Hasik, Ph.D. January 14, 2021

NO. 8. *Building Industrial Resilience with a Little Help from Our Friends,* Jerry McGinn, Ph.D. June 5, 2021

NO. 9. *Achieving Defense Exportability,* Frank Kenlon. August 16, 2021

NO. 10. *Implementing Responsible AI: Proposed Framework for Data Licensing,* Major Andrew Bowne and Benjamin McMartin. April 29, 2022

NO. 11. *"Don't Get Ahead of the Data": Artificial Intelligence and Predictive Maintenance in DoD,* Jerry McGinn, Richard Beutel, and Benjamin McMartin. June 27, 2022

NO. 12. *NASA in the Modern Space Age: A Leader Or a Follower?* Hina Kazmi, Ph.D. October 7, 2022

NO. 13. *Pathways to Defense Budget Reform,* Eric Lofgren. November 1, 2022

NO. 14. *Importance of Customer Assets in Federal Mergers and Acquisitions,* Ju-Yeon Lee, Ph.D., Brett W. Josephson, Ph.D., Shuai Yan, Ph.D. December 5, 2022

NO 15. *Improving Economy and Efficiency in Federal Contracting: Presidential Use of the Federal Property and Administrative Services Act to Direct Procurement Policy,* Emily W. Murphy. December 12, 2022

NO 16. *Back to the Future? Second Sourcing in Defense Acquisitions*, *Olivia Letts,* Jerry McGinn, Ph.D., Richard Beutel. July 12, 2023

NO 17. *Small Disadvantage Business Goals: The Effects of Recent Administrative Changes,* Emily W. Murphy. August 23, 2023

NO. 18 *Effective Competition and Market Concentration Trends in the Department of Defense Contractor Base,* Edward Hyatt. November 2, 2023

NO. 19 *How Not to Alienate Business Partners: A Framework for Addressing Factors Impacting Retention of Defense Contractors,* Moshe Schwartz and Michelle V. J. Johnson. November 2, 2023

NO. 20 *The Pote*ntial Effects of Raising the Micro-Purchase Threshold, Emily W. Murphy, John G. (Jerry) McGinn, and Richard Beutel. June 4, 2024

NO. 21 *PPBE Impact on Technology Transition, Findings and Recommendations,* Jeff Kojac, Olivia Letts, Edward Hyatt, Ph.D., and Jerry McGinn, Ph.D. September 5, 2024