QUESTIONS IN PUBLIC ADMINISTRATION

by

Mary Clare Peate
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Economics

Committee:

____ Director

____

____

____ Department Chairperson

____ Program Director

____ Dean, College of Humanities and Social Sciences

Date: ___________________________________ Spring Semester 2020
George Mason University
Fairfax, VA

Questions in Public Administration

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Mary Clare Peate
Master of Public Policy
University of Chicago, 2008

Director: Garett Jones, Associate Professor
Economics

Spring Semester 2020
George Mason University
Fairfax, VA

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

QUESTIONS IN PUBLIC ADMINISTRATION

Mary Clare Peate, Ph.D.

George Mason University, 2020

Dissertation Director: Dr. Garett Jones

My dissertation looks at three different public administration challenges and explores the effect that bias, tradeoff of efficiency and accuracy, and unintended consequences have on decision making.

The first public administration challenge is bias in administering public services. Using disability appeals at the Social Security Administration, I find that traits such as political ideology, gender, and age of administrative law judges robustly predict the outcomes of disability appeals at the Social Security Administration. To my knowledge, this is the first research to estimate partisan effects on disability decisions and the first to incorporate the third common political self-identification of Independent in an analysis of quasi-judicial decision making. Democratic judges are most lenient, Independents are least lenient; male judges are more lenient than female judges. My

model predicts that a female Independent judge awards disability benefits at 10.1 percentage points less often than a male Democratic judge.

The second chapter analyzes the tradeoff between accuracy and efficiency in public administration. How do time constraints shape approval rates of disability appeals at Social Security Administration? The random assignment of Social Security disability cases to U.S. administrative law judges—and the requirement that they adjudicate 500 cases per year to remain in good standing—provide an opportunity to offer some answers to this important question. The quota on the number of cases an administrative law judge must adjudicate in a year robustly predicts outcomes for a small subset of administrative law judges, 6% of my sample, who appear to be time constrained. These judges are 2.1% more likely to approve appeals cases throughout the entire year and 3.7% more likely to approve appeals cases in the final month of their fiscal year.

The third chapter tackles the final challenge: unintended consequences when awarding government contracts. Federal procurement has a goal of awarding 23% of all federal contract dollars to small businesses. I look at very small firms participating in the federal procurement economy and find that they differ from small firms in the regular economy in what types of industries they participate in. Additionally, small firms' success by contract dollars obligated is strongly positively correlated with how many industries they participate in: Winning a contract in an additional 4-digit NAICS code predicts a 48% increase in contract dollars. I present evidence that participating in multiple NAICS codes is a proxy for understanding the procurement market's regulation, processes, and networks rather than authentic diversification. This suggests that success in the procurement market for small firms has less to do with expertise in an industry and more to do with specialization in federal procurement.

1. PARTY AFFILIATION AND GENDER AFFECT JUDICIAL DISABILITY APPEALS CASES

1.1 Introduction

In fiscal year 2016, nearly 2.4 million Americans filed applications with the Social Security Administration (SSA) for disability benefits. Roughly 1.6 million (67%) of these applications were rejected.[1] For these rejected applications, almost 700,000 applicants appealed the SSA's decision.

Every year, SSA's administrative law judges (ALJs) consider appeals denying requests for disability benefits. Although roughly 45% of all such appeals are approved[2], approval rates vary considerably by judge. For example, in 2017, Administrative Law Judge Walters adjudicated 633 disability appeals cases and awarded benefits in 472, or 75%, fully favorably. In the same year, Administrative Law Judge Batik heard 396 cases and decided only 39, or 9.8%, fully favorably. The other roughly 1,600 ALJs had approval rates somewhere in between. Figure 1.1 shows the distribution of approval rates in 2017.[3]

---

[1] Social Security Administration (n.d.); author's calculation.

[2] This approval rate does not include dismissal data; with dismissal data the average approval rate is approximately 37%. All further calculations below include dismissal data.

[3] Social Security Administration (2020); author's calculation.

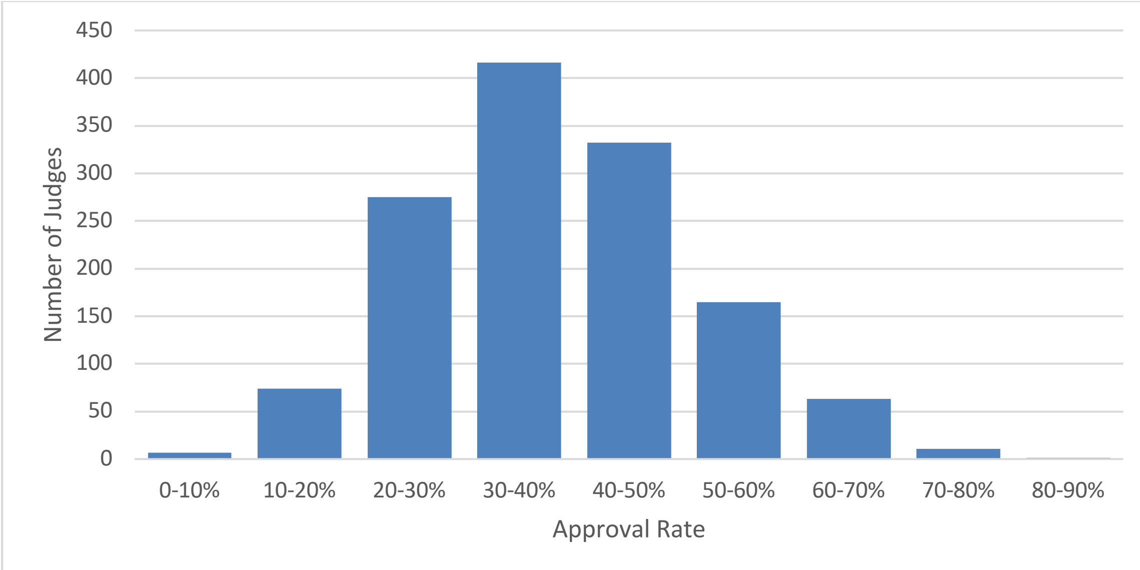


Figure 1.1. Administrative Law Judge Disability Appeal Approval Rates (2017), author's calculation.

Some variation in approval rates can be explained by geographic heterogeneity, but its persistence at the office level where disability appeals cases are randomly assigned[4] indicates judges also exercise their discretion. In fact, Judge Walters and Judge Batik mentioned above are in the same Dallas office. Figure 1.2 displays the approval rates of all of the judges in that office.

[4] Appeals cases are randomly assigned at the office level, according to the ALJ's HALLEX (Hearings, Appeals, and Litigation Law) manual I-2-1-55 (SSA, 2019). I also confirmed that cases are randomly assigned through a conversation with an administrative law judge.

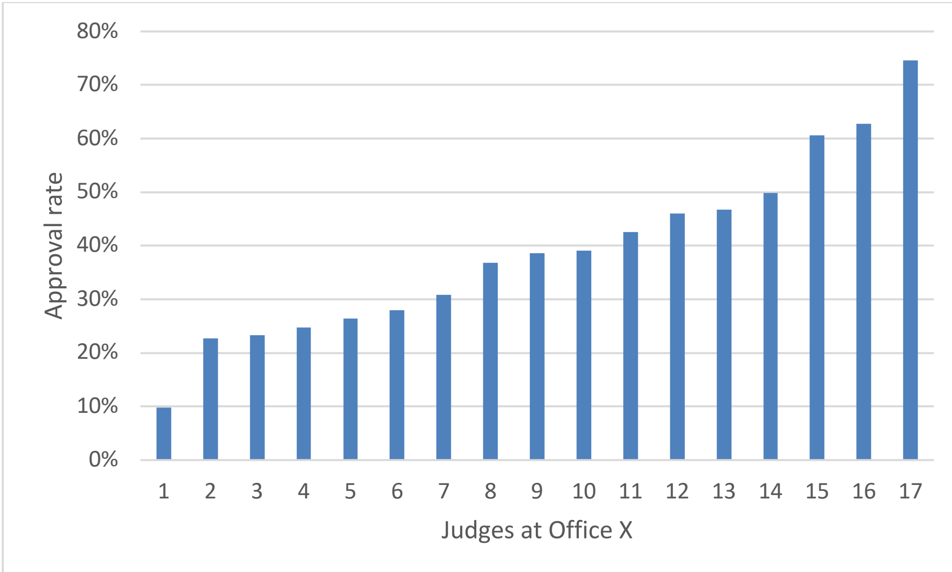


Figure 1.2. Approval Rates by Administrative Law Judge, Dallas office (2017).[5]

Figure 1.3 replicates Figure 1.2 but controls for office-specific differences. The approval rate variation amongst ALJs shrinks considerably but still remains.

[5] Social Security Administration (2020); author's calculation.

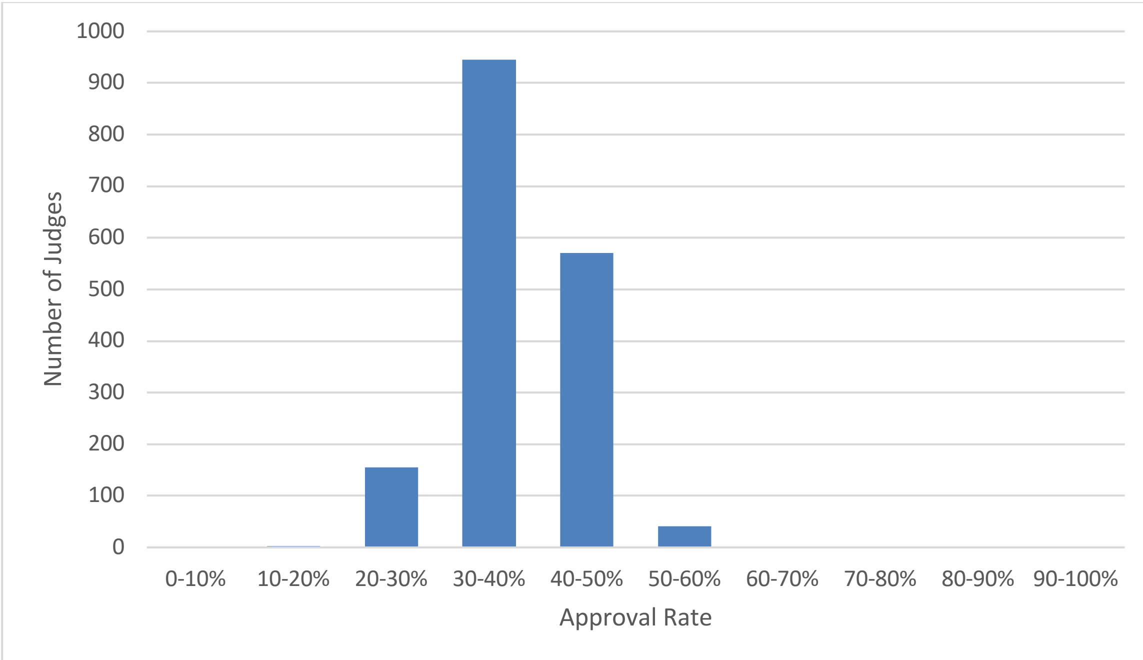


Figure 1.3. Administrative Law Judge Disability Appeal Approval Rates (2017), Controlling for Office.

Given that disability appeals cases are randomly assigned at the office level, I ask whether any of this remaining variation amongst ALJs' appeals approval rates is correlated with political ideology, and find that judges who identify as Democrats are on average the most lenient judges, followed by Republicans and then Independents.

Specifically, in a fractional response logistic regression of political affiliation, gender, and age on a judge's disability appeals approval rate in fiscal years 2016-2018, controlling for office and fiscal year, I find that Democratic judges have an approval rate 4.2 percentage points higher than Republican judges and 5.0 percentage points higher than Independent judges. Political ideology is not the only characteristic that matters: Male judges have approval rates 2.7 percentage points higher than female judges.

Gender and political ideology further interact to produce additional statistically significant differences in approval rates. Finally, age is also positively correlated with disability appeals acceptance rates. Combining political ideology, gender, and age can change the disability appeals acceptance rate by up to 14.2 percentage points: My model predicts that a 50-year-old female Independent judge has an approval rate **14.2 percentage points** lower than a 67-year-old male Democratic judge.

The rest of this chapter is organized as follows: Section 2 is a literature review, Section 3 presents the empirical strategy, Section 4 summarizes data used, Section 5 explains the empirical specification, Section 6 presents results, and Section 7 concludes.

1.2 Literature Review

There is significant prior literature analyzing factors affecting judicial decision making.[6,7] However, most of this literature does not specifically estimate factors affecting administrative law judges' (ALJ) decision making.

One of the first studies of ALJ decision making was conducted by Mashaw et al. (1978). Using a survey of 180 ALJs, they present findings about the relationship between ALJs' attitudes about disability processes and their approval rates. For example, if a judge agrees with the statement "disability standards are too stringent," their approval rate was 18% higher than those who did not. Interestingly, the survey asks respondents

---

[6] For a nice summary of the literature, see *The Behavior of Federal Judges* by Epstein et al. (2013).
[7] Analyzing judicial behavior is so common that France wants to ban the practice (Artificial Lawyer, 2019)!

if they identify as "liberal" and find no correlation between their responses to this question and approval rates.

More recently, Nakosteen and Zimmer (2014) conducted the most prominent empirical analysis of factors affecting judicial decision making on disability appeals cases. While controlling for personal characteristics that include gender and judicial experience, they ask whether exogenous economic and political factors affect judicial decision making and find that a judge's approval rate is positively correlated with her state's unemployment rate and a governor's party identification of Democrat. They also find that a judge's years of judicial experience is positively correlated with her disability appeals approval rate.

However, to my knowledge, there has been no direct analysis of the relationship between party affiliation and administrative law judges' disability appeals approval rates. Research on party affiliation and other types of judicial decision making, on the other hand, is fairly robust. Most U.S. federal and state judges are either appointed or elected. In either case, the judge's political affiliation is strongly suspected, making it easier to analyze the relationship between party affiliation and legal decisions.

For example, Sunstein et al. (2006) analyze federal district courts of appeals decisions and find that on cases involving controversial issues such as affirmative action, sex discrimination, and 11th amendment,[8] judges appointed by Republican presidents

---

[8] Full list: affirmative action, National Environmental Policy Act, 11th amendment, National Labor Relations Board, sex discrimination, disability, sexual harassment, campaign finance, piercing the corporate veil,

decide cases differently than those appointed by Democratic presidents. Perhaps more surprisingly, Sunstein et al. found substantial dampening effects across party lines when the panel of judges included judges appointed by both Republican and Democratic presidents. However, this dampening effect did not appear to apply to cases involving abortion or capital punishment.

Furthermore, Sunstein et al. do not find variation in decision making by political ideology on cases involving criminal appeals, federalism commerce clause, takings, punitive damages, and standing. They speculate that there is less room for interpretation for these types of cases.

Given this current research, it is unclear whether party affiliation will affect SSA administrative law judges' decisions for several reasons. First, assume that administrative law judges' behavior will be consistent with Sunstein et al.'s findings: Will disability appeals cases be more like controversial affirmative action cases where there are substantial ideological differences in decisions, or like criminal appeals cases where there are not? Furthermore, Sunstein et al. and others do not disaggregate judicial decision making by Independent judges because they do not have that data. If a judge identifies as Independent, will he or she behave like Republicans, Democrats, or neither?

---

environmental regulation, obscenity, Title VII, racial segregation, Federal Communications Commission, contracts clause violations, 1st amendment challenges to commercial advertising restrictions, abortion, and capital punishment (Sunstein et al., 2006, p. 24).

However, there is no reason to assume that judges' behavior at the federal court of appeals will predict ALJ behavior on disability appeals. Administrative law judges and their appeals are very different from federal district judges and their appeals. First, the types of people who select into an administrative law judge career are very different from those who select into federal and state judge positions. Just 6% of district court judges were either a magistrate, bankruptcy, or administrative law judge before appointment (Epstein et al., 2013, p. 341),[9] suggesting the kinds of individuals who become district court judges are different from those who become administrative law judges.

Second, federal judges are presidentially appointed, senate confirmed, and tenured for life. Similarly, most state judges are either appointed by the governor or elected. In these instances, their political partisanship is a feature, not a bug. Not so with administrative law judges. Administrative law judges are ultimately hired by a federal agency, but they must first apply for the job through the Office of Personnel Management (OPM) which screens and selects the top three applicants for a position. The agency director can then hire one of these three applicants.[10] In fact, the hiring process of administrative law judges is very similar to that of a typical federal employee who is also screened and selected by OPM based on criteria before the hiring agency

---

[9] Epstein et al. (2013) do not further disaggregate between magistrate, bankruptcy, and administrative law judges, but their language suggests administrative law judges are the least common of the three.
[10] The process is different because of the recent Supreme Court ruling *Lucia et al. v. Securities and Exchange Commission* (2018), but my dataset has judges selected under these old procedures.

can review the OPM-approved applicants. To my knowledge, there is no research on political affiliation affecting government bureaucrats' decision making.

Furthermore, disability appeals cases are low profile and low stakes in contrast with many of the cases decided at the federal or state court or court of appeals. Even uncontroversial cases at the federal court of appeals garner much more national attention than a disability appeals case. If I apply Friedman and Friedman's (1962) theory of taste-based discrimination to judicial decision making, lower profile cases with lower stakes should afford more discretion and more bias, not less, predicting political affiliation will affect outcomes. There is very little research on judicial decision making and political ideology for non-Article III judges, such as magistrate, bankruptcy, or immigration judges.

This chapter will contribute to the literature by estimating whether political affiliation affects decision-making in disability appeals in which the stakes are relatively lower and the ALJs have a greater level of discretion. Additionally, this is the first research to my knowledge that will include and disaggregate by *Independent* political affiliation when analyzing results of political affiliation on judicial decision making.

1.3 Empirical Strategy

I follow Nakosteen and Zimmer's (2014) empirical strategy, albeit simplified in three ways to focus on my particular research question. Their model, based on Marinescu's (2011) model used to analyze whether economic conditions affect judges' decision making in employment dismissal cases, is summarized as follows: Suppose M_i^*

represents $judge_i$'s perceived merits of an appeal, and M_i^O represents $judge_i$'s objective standard of an appeal in the absence of subjective factors, denoted M_i^S. If $M_i^S = 0$, then a judge approves the appeal solely according to her objective standard, whenever M_i^* > M_i^O. If $M_i^S \neq 0$, then a judge approves a disability appeal if M_i^* > M_i^O+ M_i^S.

I estimate M_i^s; first I divide M_i^s into two general components of subjectivity: a judge's individual characteristics, denoted X_i, and the judge's environment, denoted Z_i. The model is as follows: $M_i^S = B_0 + B_1X_i + B_1Z_i$. For a judge's individual characteristics, X_i, I include three variables: political affiliation, gender, and age.

While the focus of my research is on the relationship between a judge's party affiliation and her approval rates, other variables may be correlated with political affiliation and omitting them would bias results. This is why I include gender. While there is ample evidence that gender affects judicial decision making, research is not conclusive and there is far less research about how or if a judge's gender interacts with party affiliation. For example, Gruhl et al. (1981) found that female trial judges were less likely to find defendants guilty but more likely to send convicted defendants to jail and more likely to apply harsher sentences. On the other hand, Schrag et al. (2009) find that, after controlling for past professional experience, female federal immigration judges were 30% more likely to grant asylum to asylum seekers than male immigration judges (p. 346). While there is ample evidence that gender affects judicial decision making, there is far less research about how or if gender interacts with party affiliation.

Additionally, including gender in my analysis allows me to perform a replication test for Nakosteen and Zimmer's (2014) surprising result that gender had no effect on approval rates for social security disability cases.

Finally, I include age as one of a judge's characteristics, X_i, for two reasons. First, age is correlated with gender in my dataset: There is a higher proportion of younger female judges. Therefore, omitting age could overstate gender's effect.

Table 1.1 shows 2017 age distribution by gender. Females skew younger in this dataset.

Table 1.1. 2017 Administrative Law Judge Age Distribution by Gender

Age range	# of female judges	% of female judges	# of male judges	% of male judges
30-39	2	0.7%	3	0.6%
40-49	56	18.7%	80	14.8%
50-59	133	44.3%	150	27.8%
60-69	93	31.0%	192	35.6%
70-79	15	5.0%	105	19.5%
80-89	1	0.3%	9	1.7%
Total	**300**	**100%**	**539**	100%

Second, including age could also clarify Nakosteen and Zimmer (2014)'s finding that judicial experience is positively correlated with approval rates. Age and judicial experience are positively correlated, so if I find that age has no statistically significant effect or a negative effect on approval rates, it would strengthen the case that Nakosteen and Zimmer's (2014) finding indicates judicial experience is positively correlated with approval rates. However, if I find that age is positively correlated with approval rates, then perhaps their finding was estimating mere chronological effects of age rather than judicial experience effects.[11]

The second component of subjectivity in M_i^s is external factors, denoted Z_i. I simplified Nakosteen and Zimmer's (2014) model by only including office and fiscal year individual and interactive effects for Z_i. Nakosteen and Zimmer (2014) include a variety of state-specific economic and political factors such as a state's unemployment rate or the political affiliation of its governor as well as the office-specific factor of number of backlogged cases. However, the relationship I sought to measure is not between *specific* external environmental factors and approval rates but the relationship between a judge's political affiliation and her approval rates. Therefore, I do not need to know the strength or sign of the environments' effects on approval rates, but I do need to control for them. Hence, I include office level and fiscal year in my model, effectively controlling for any external environmental effects at the state or office level.

I also include the number of cases adjudicated by a judge although Nakosteen and Zimmer (2014) find no correlation between this metric and approval rates. Figure 1.4 plots approval rates and number of decisions rendered by all judges adjudicating

[11] And yes, obviously it would be better to include both age and judicial experience to settle the question once and for all, but no one would give me this data.

more than 300 cases in 2016-2018. I excluded those judges with a very low number of decisions because they usually have not worked for the full year and thus, it is difficult to extrapolate from their behavior. A simple linear regression implies a very slight positive relationship between number of adjudicated cases and approval rates.

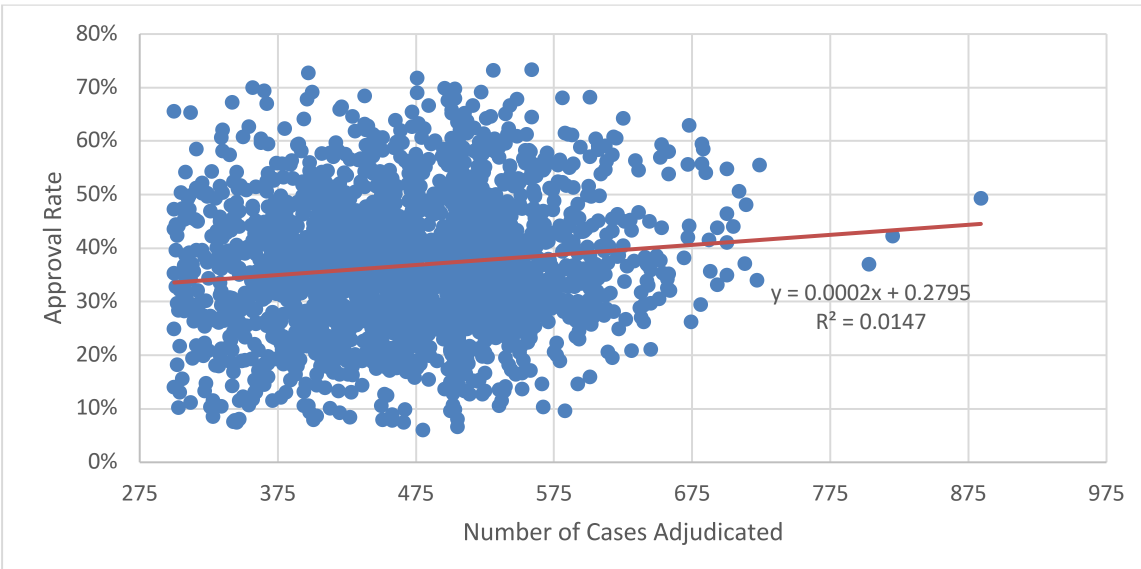


Figure 1.4. Administrative Law Judge Approval Rate by Adjudicated Cases (My Dataset Excluding Those Below 300 Cases).

Nakosteen and Zimmer (2014) claim that if judges are issuing more approvals, demand for their services might increase. Given that cases are assigned randomly at the office level, I do not believe this should be an issue. A simpler explanation is that approving an appeal requires far less work than rejecting one. The reason for this is straightforward: If an appeal is denied, the appellant can appeal this claim again at the federal court of appeals. The decision must withstand scrutiny at the next level and therefore documentation must be more rigorous. On the other hand, an appeal approval receives far less attention. As Marchand and Warshawsky (2015) state. “If the ALJ then awards disability benefits, the decision is final because the government does not appeal it” (p. 4). The Office of Inspector General of SSA agrees:

> Denied claims typically have a longer hearing, longer time writing the case, and longer ALJ instructions to decision writers. Claimants can appeal denials, so the written decision has to be inclusive of all factors raised by the claimant or during the hearing to withstand scrutiny on appeal. (Office of the Inspector General, Social Security Administration, 2017, p. 8)

While it makes sense to confirm that higher approval rates are positively correlated with higher caseloads by a simple correlated scatterplot, Nakosteen and Zimmer’s (2014) conclusion that this will attract demand for their services does not follow. According to HALLEX, cases are randomly assigned, making it difficult for lawyers to demand judicial services (SSA, 2019).

So yes, a higher number of cases decided could be correlated with higher approval rates, but the correlation is not particularly interesting because the causality appears to run in the opposite direction: Higher approval rates could cause judges to decide more cases.

Finally, I depart from Nakosteen and Zimmer’s (2014) model in one more decision. They include office caseload data but find there is no correlation. I chose not to include caseload data because I disagree with their strategy for including this variable.

1.4 Data and Variable Descriptions

There were three main data sources in this chapter: SSA's publicly available disposition data set, Aristotle's voter lists, and Martindale's lawyer directory database. The primary dataset comes from the Social Security Administration's Administrative Law Judge (ALJ) Disposition Data for fiscal years 2016-2018 (SSA, 2020). Federal fiscal years begin October 1st and end September 30th. This dataset includes the judge's name, office, total dispositions at her current office and at any other offices, decisions, awards, denials, and the number of awards that were fully or partially favorable. There are 184 offices, 10 regions, and approximately 1,600 ALJs deciding nearly 700,000 cases every year.

I obtained ALJ voter registration records from Aristotle's voter lists (Aristotle.com),[12] an online voter identification database. To match administrative law judge data from SSA's disposition database to Aristotle's database, I followed Langbert et al.'s (2016) decision rules in their paper measuring the degree of political diversity in university humanities departments. In particular, in order for them to link a professor with a registration record in Aristotle, the person's name, geographic location, and age needed to match.

---

[12] ALJs registered in Pennsylvania and Arizona were excluded from the study because I was not allowed to access them through Aristotle.

To match on name, an ALJ's first and last names as well as middle initial had to match the Aristotle record.[13] To match on location, the Aristotle record had to be less than 100 miles from an ALJ's office location. And finally, to match on age, an ALJ's age had to match the Aristotle record's age. To approximate an ALJ's age, I used the Martindale lawyer directory database (Martindale.com). Lawyers self-report personal information such as university and law school graduation year, and year they passed the bar. I constructed an approximate age based on estimating their birth year from their college or law school graduation year; if Aristotle's record was within 5 years of a university graduation-constructed age or 10 years of bar passage/law school graduation-constructed age, I matched and used Aristotle's age rather than Martindale's age in my model. I was more stringent with the university-constructed age because there is less variation when individuals attend and graduate from college.

If more than one person fit these criteria in Aristotle, I did not use the record. I confined this search to ALJs employed in 2018 only since confirming voter registration data in prior years would be difficult and less reliable. Additionally, I only included ALJs designated as Republican, Democratic, or Independent because there were too few identified by other labels. In total, I dropped four individuals because of uncommon party identification: one Green party, two Libertarians, and one identified as "other."

[13] I used FedPay.org to obtain the middle initial when it was not included in the ALJ disposition dataset.

The Aristotle database also records gender, so I confirmed an individual's gender with that record.

I excluded individuals who had heard fewer than 50 cases in a year. Functionally, this means I excluded newly hired ALJs who started late in a given fiscal year and ALJs retiring at the start of a fiscal year. For simplicity, I assigned each ALJ to one office per fiscal year and excluded ALJs' disposition data from other offices within a given year. This means that I lost some fidelity because I dropped observations of one office when a person transferred midyear. Finally, in order to be included in my dataset, ALJs had to have 3 full years of disposition data. I was left with 839 administrative law judges with 3 years of decision data. Table 1.2 reports summary statistics of judge characteristics: 64% of judges are male and 36% are female; 46% are Democratic, 31% are Republican, and 22% are Independent.

Table 1.2. Summary Statistics of Administrative Law Judge Characteristics

Gender	Democrat	Republican	Independent	Total
Female	183	61	56	300
Male	203	202	134	539
Total	386	263	190	839

1.5 Specification

Similar to Papke and Wooldridge (1996), I use a fractional response logistic model because the dependent variable is a continuous variable bounded from 0 to 1. Although I have 3 years of data for 839 judges, I pooled the data because I wanted to estimate gender and political affiliation which, in my dataset, are time-invariant characteristics.

Below is the model:

$$\mathrm{Ln}\left[\frac{P\_APP}{1-APP}\right] = \alpha + b_1 gender + b_2\, registered\ political\ party + b_3 gender * registered\ political\ party + b_4 age + b_5 number\ of\ cases + b_6 fiscalyear + b_7 office + b_8 (fiscalyear * office) + \varepsilon \quad [1]$$

Cases are assigned randomly at the office level. Therefore, omitted variables will not be correlated with characteristics of the case, although they may be correlated with the individual traits of judges I do control for.

1.6 Empirical Results

Table 1.3 presents estimates from Equation 1, the logistic model estimating the proportion of fully favorable approvals. Several personal characteristics, X_i, and external factors, Z_i, are statistically significant. First, controlling for office and fiscal year, males, Democrats, and older judges are more likely to have higher disability appeals approval rates.

Table 1.3. Proportion of Administrative Law Judge's Fully Favorable Approvals, Controlling for Office and Fiscal Year

	(1) Approval rate	(2) Approval Rate
male	0.117***	0.119***
	(4.18)	(5.78)
Republican	-0.0931*	-0.190***
	(-2.38)	(-8.13)
Independent	-0.331***	-0.224***
	(-7.05)	(-8.96)
male*Republican	-0.126**	
	(-2.69)	
male*Independent	0.149**	
	(2.68)	
age	0.0118***	0.0116***
	(11.00)	(10.84)
cases	0.000299***	0.000293***
	(3.46)	(3.37)
N	2517	2517

t statistics in parentheses
* *p*<0.05, ** *p*<0.01, *** *p*<0.001

Table 1.4. Odds Ratio of Administrative Law Judge's Fully Favorable Approval Rate, Controlling for Office and Fiscal Year

	(1)
	Odds ratio
male	1.124***
	(4.18)
Republican	0.911*
	(-2.38)
Independent	0.718***
	(-7.05)
male*Republican	0.881**
	(-2.69)
male*Independent	1.161**
	(2.68)
age	1.012***
	(11.00)
cases	1.000***
	(3.46)
N	2517

Exponentiated coefficients; t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 1.4 shows the odds ratio of administrative law judge's fully favorable approval rate, controlling for office and fiscal year. Independents are the least lenient judges, with the lowest approval rates, followed by Republicans, and then Democrats.

And because I am controlling at the office level, this effect is independent of state economic or political effects as described in Nakosteen and Zimmer (2014).

Similar to Nakosteen and Zimmer (2014), I find a statistically significant relationship between age of the judge and leniency. I added age squared as an additional parameter to test whether the age's effect on approval rates was nonlinear, but this was not statistically significant and reduced the age parameter effect, so I dropped it from the model.

However, my finding on another personal characteristic, gender, is not similar to Nakosteen and Zimmer (2014). Here, I find that males are more lenient than females and the effect is statistically significant.

Reporting the marginal effects of binary parameters is fairly straightforward. Male judges have an approval rate **2.7 percentage points** higher than females. Republican judges have approval rates **4.3 percentage points** lower than Democratic judges and Independents have approval rates **5.0 percentage points** lower than Democratic judges. There is no statistically significant difference between Republican and Independent approval rates.

The above results do not include any interactive effects between gender and political affiliation. If I include interactive effects, a female Independent judge has an approval rate **10.1 percentage points** lower than a male Democratic judge. The difference between a Democratic male judge and an Independent female judge is statistically significant (Table 1.5, Figure 1.5).

Table 1.5. Predicted Margins by Gender and Political Affiliation of Administrative Law Judge's Fully Favorable Approval Rate, Controlling for Office and Fiscal Year

	Margin (Delta-method)	Std. Err.	z	P>|z|
gender# voteridstatad0r~2				
Female # Democratic	.3767364	.0046138	81.65	0.000
Female # Republican	.3556168	.0074699	47.61	0.000
Female # Independent	.3043027	.008617	35.31	0.000
Male#Democratic	.4039552	.0042338	95.41	0.000
Male#Republican	.353597	.0044384	79.67	0.000
Male#Independent	.362085	.0052106	69.49	0.000

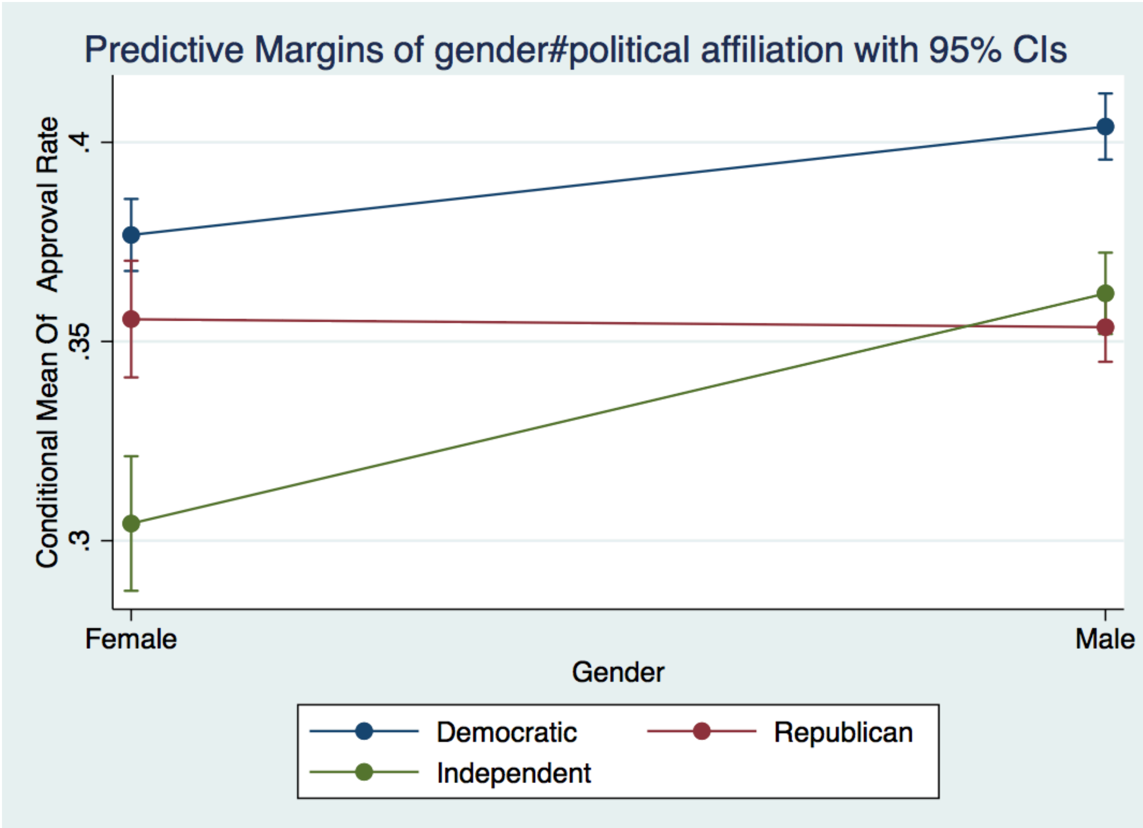


Figure 1.5. Predictive Margins by Gender and Political Affiliation with 95% Confidence Intervals.

To factor age into the analysis, I report the difference between a judge at the 20th and 80th percentile of the age distribution in the dataset. For logistic regressions, it is easiest to compare two different profiles. The difference in approval rates between a 50-year-old Independent female judge, the 20th percentile of the age distribution in my dataset, versus a 67-year-old Democratic male judge, at the 80th percentile of the age distribution of my dataset, is **14.2 percentage points.**

For Z_i, although my z parameters were much more limited than Nakosteen and Zimmer's, I also had a discrepancy with their finding regarding the relationship between number of cases decided and approval rates. I find that number of cases is also statistically significant and positively correlated with approval rates. This finding supports the claim that approving an appeal takes less time than denying one.

Nakosteen and Zimmer (2014) found a positive but statistically insignificant relationship between cases decided and approval rates. There are two potential reasons for the different results for this variable. First, they estimated the effect using two stages because they worried about endogeneity. I disagree that endogeneity is an issue because cases are randomly assigned at the office level and therefore estimated it directly in my model.

Second, their methodology for considering records could differ from mine. I dropped a judge record any time a judge decided fewer than 50 cases in a fiscal year. If I instead kept this data in my dataset, it could potentially bias results. Judges hearing fewer than 50 cases are not working a full fiscal year. If they only decide, say, 10 cases, their approval rate may be different from their true approval rate if they were to hear and decide cases for an entire year. Although 50 cases is an arbitrary number, I chose the cutoff of number of cases to be large enough that it would likely represent their natural approval rate. Nakosteen and Zimmer (2014) do not mention number of cases as a deciding factor in keeping a record.

1.6.1 Robustness Checks

Additionally, I performed two different robustness checks. First, I varied the regression method and used a linear regression model instead of a fractional response logistic method (Table 1.6). I also included a beta regression in Appendix A; the results do not vary too much.

Table 1.6. Linear Regression Robustness Test

	(1)
male	0.0274***
	(3.83)
Republican	-0.0216*
	(-2.21)
Independent	-0.0728***
	(-6.52)
Male*Republican	-0.0284*
	(-2.41)
Male*Independent	0.0305*
	(2.26)
age	0.00268***
	(9.89)
cases	0.0000693**
	(3.17)
N	2517

t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Second, I changed the response variable in two different ways. First, judges have the option of deciding cases fully favorably or partially favorably for the disability appellant. In all previous analysis, I had estimated results according to judges deciding a case fully favorably. Rather than estimating the proportion of cases decided fully favorably, I used the proportion of appeals cases decided fully or partially favorably. Results remain similar but weaker; Table 1.7 reports. Republican is no longer statistically significant different from Democrat. The interactive dummy variable of Independent males becomes statistically insignificant.

Table 1.7. Response Variable Is Proportion of Partial or Fully Favorable Approvals, Controlling for Office and Year

	(1) Approval rate	(2) Approval rate
male	-0.133***	-0.143***
	(-4.04)	(-6.16)
Republican	-0.0228	0.0224
	(-0.53)	(0.76)
Independent	0.134**	0.0728**
	(3.18)	(2.59)
male*Republican	0.0563	
	(1.07)	
male*Independent	-0.0930	
	(-1.74)	
age	-0.00502***	-0.00492***
	(-4.14)	(-4.07)
cases	-0.000103	-0.000101
	(-0.99)	(-0.98)
N	2517	2517

t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Second, I created a new response variable of judicial approval rate by restricting focus to those cases heard by a judge. An appeal can either be dismissed or heard. If a case is heard, it is either denied in full, approved in full, or partially approved. If I exclude cases that are dismissed and focus only on cases that are heard as reported below, the effects on male, Republican, and Independent are stronger. This could be because there is less discretion when dismissing a case. I also varied controls within this method and tried controlling by office only, state only, and state and fiscal year.

Table 1.8 reports results. They are similar. The only real difference is that when using state instead of office, once again, the interactive dummy of Independent and male fails to be statistically significant. It was weakly statistically significant at the 5% level in my original model.

Table 1.8. Response Variable Is Percent of Cases Decided Fully Favorably for Only Heard Cases

Variable	(1) (office*yr)	(2) (office*yr)	(3) (state*yr)	(4) (state*yr)
Male	0.154***	0.148***	0.141***	0.123***
	(4.83)	(6.33)	(4.50)	(5.16)
Republican	-0.0948*	-0.215***	-0.112*	-0.226***
	(-2.14)	(-7.95)	(-2.47)	(-8.09)
Independent	-0.370***	-0.265***	-0.324***	-0.250***
	(-7.11)	(-9.30)	(-6.27)	(-8.29)
male*Republican	-0.160**		-0.155**	
	(-3.00)		(-2.85)	
male*Independent	0.147*		0.0989	
	(2.35)		(1.58)	
age	0.0131***	0.0130***	0.0135***	0.0133***
	(10.72)	(10.58)	(11.28)	(11.18)
cases	0.000215*	0.000209*	0.000272**	0.000267*
	(2.18)	(2.10)	(2.63)	(2.57)
N	2517	2517	2517	2517

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

1.7 Conclusion

Gender, political affiliation, and age can potentially impact how public programs are administered, including specifically who receives benefits. Although the individual effects of each trait are not large, the cumulative effect on a judge's disability appeals approval rate can be striking: A 67-year-old male Democratic judge has an approval rate 14.2 percentage points higher than a 50-year-old female Independent judge. Furthermore, each administrative law judge decides roughly 450 cases per year, and there are 1,600 judges adjudicating disability appeals. Minor bias can affect tens of thousands of appeals cases per year.

Being accepted into a disability program can affect outcomes. French and Song (2014) report that effective random assignment into the disability program because of a more lenient ALJ reduced the recipient's labor supply 3 three years later by 26% percentage points. Maestas et al. (2013) had a similar finding when looking at those who were accepted into the disability program because a lenient examiner reviewed their application.

And this is but one example of bias in but one program: Many agencies are reviewing applications and petitions to receive some type of benefit. Legislative and administrative agencies designate criteria for accepting applicants, much like SSA does for disability insurance, with the understanding that these criteria will be applied in a uniform and consistent manner. Human bias undermines these goals, potentially turning the administration of important government programs into a quasi-lottery. Understanding that human bias is part of this process could inform changes to how we structure government programs moving forward, including developing techniques to minimize the impact of bias.

Furthermore, because administrative law judges are the focus of this analysis, some of the findings cross over to the judicial branch as well. Most of the judicial decision-making literature fails to report on the behavior of a judge with a self-identified Independent political affiliation, despite the fact that 42% of Americans identify as

Independent (Jones, 2018). This chapter suggests there may be a difference in how Republicans and Independent judges decide disability appeals cases. When possible, future research on judicial decision making and political affiliation should distinguish between self-identified Republican and Independent judges.

2. WAVING SOME THROUGH: DISABILITY APPEALS CASE QUOTAS CAN AFFECT APPROVAL RATES

2.1 Introduction

In 2007, the backlog of disability appeals at the Social Security Administration (SSA) reached 750,000; the average processing time for an appeal was roughly 500 days, almost double the wait time in 2000 (SSA, 2016, p. 3), when then-SSA Commissioner Michael J. Astrue testified before Congress that one of his priorities was to reduce this backlog (SSA, 2007). Several months later, SSA issued a case adjudication goal, or quota, for their administrative law judges (ALJ) handling disability appeals: ALJs' performance would be evaluated in part by whether they adjudicated at least 500 cases annually (SSA, 2012). SSA administrative law judges, through the Association of Administrative Law Judges (AALJ), sued SSA, claiming a quota would jeopardize their judicial independence.[14] They argued that because an approved disability claim is not subject to further appeal, combined with the fact that approving an appeal takes less time than

[14] I tried to discover what penalties or rewards were associated with not meeting/meeting the quota, but I never received an answer. SSA maintains it is a "goal" not a quota, and AALJ, in their lawsuit asserted "the [Social Security] Administration has taken formal and informal disciplinary measures to enforce it, so that it is in effect an enforceable and enforced quota" (*Association of Administrative Law Judges vs. Colvin*, 2001, p. 2).

denying one, together implied that a quota would incentivize judges to approve more appeals. The SSA's Office of Inspector General, in a 2017 report, supports this assertion:

> Denied claims typically have a longer hearing, longer time writing the case, and longer ALJ instructions to decision writers. Claimants can appeal denials, so the written decision has to be inclusive of all factors raised by the claimant or during the hearing to withstand scrutiny on appeal. (Office of Inspector General, Social Security Administration, 2017, p. 8)

The AALJ lost their case in 2015.

Were the AALJ's claims correct? Does a case adjudication quota indeed affect disability appeals approval rates? Figure 2.1 shows the number of judges who adjudicate a particular number of cases each year for fiscal years 2016-2018. It seems clear that some judges' *number* of cases adjudicated is affected by the quota, given the unusually large increase in the number of judges who adjudicated just over 500 cases per year. There is a clear difference between the number of judges who adjudicated 499 cases each year (5 judges) and the number of judges who adjudicated 500 cases each year (24 judges). Table 2.1 summarizes Figure 2.1.

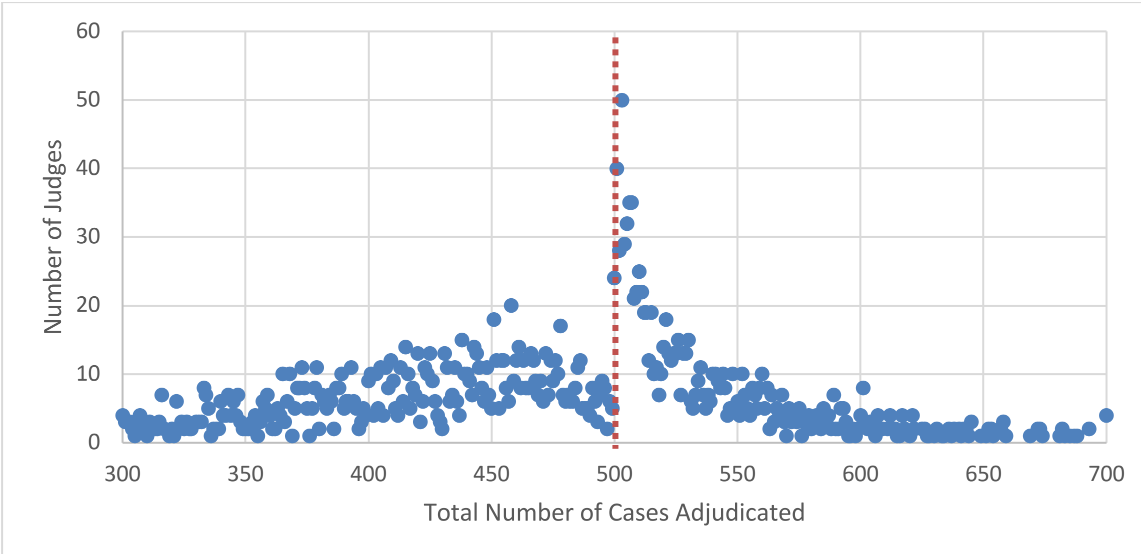


Figure 2.1. Frequency of Cases Adjudicated by Administrative Law Judges by Year (2016-2018). Data source: SSA (2020); author's calculations.

Table 2.1. Number and Percentage of Administrative Law Judges by Cases Adjudicated per Year (2016-2018)

Number of Cases Adjudicated	Number of Judges	Percentage of Judges
300-400	421	17.7%
400-410	93	3.9%
411-420	83	3.5%
421-430	67	2.8%
431-440	93	3.9%
441-450	91	3.8%
451-460	114	4.8%
461-470	100	4.2%
471-480	99	4.2%
481-490	69	2.9%
491-500	78	3.3%
501-510	**317**	**13.3%**
511-520	**143**	**6.0%**
521-530	**132**	**5.6%**
531-540	73	3.1%
541-550	75	3.2%
551-560	68	2.9%
561-570	45	1.9%
571-580	33	1.4%
581-590	35	1.5%
591-600	26	1.1%
greater than 600	122	5.1%

Note. Data source SSA (2020); author's calculations.

While it is clear from the discontinuity at 500 cases that the quota affected the number of cases adjudicated for a subset of judges, it is less obvious that the quota affected their disability appeals approval rate. In fact, Figure 2.2, which shows average approval rates by number of adjudicated cases, seems to suggest there is no change in approval rates at the cutoff of 500 adjudicated cases.

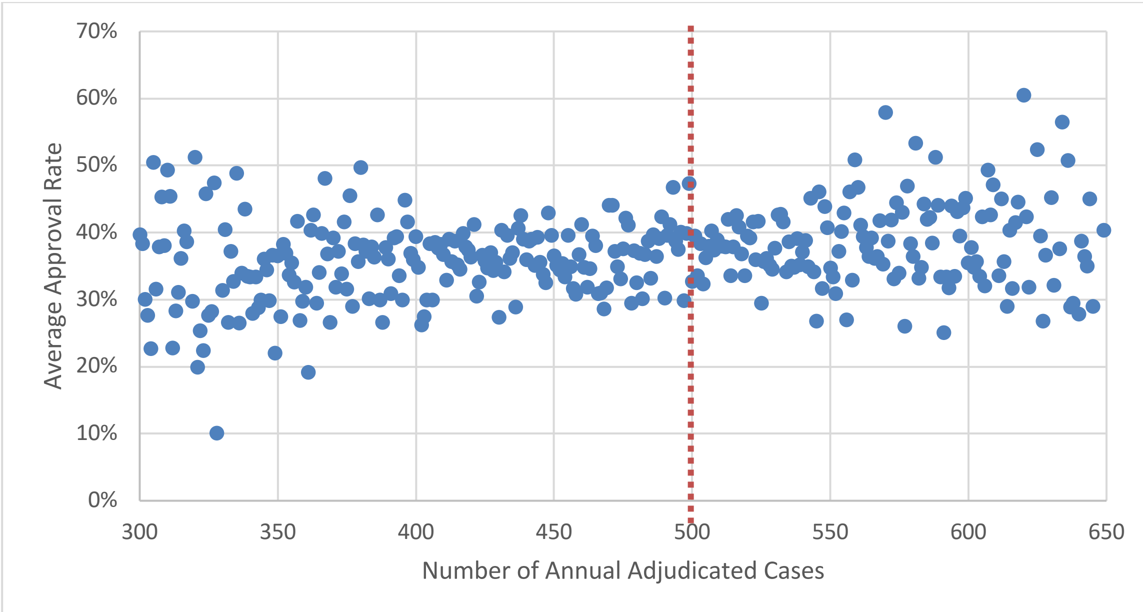


Figure 2.2. Average Approval Rate by Number of Cases Adjudicated by Administrative Law Judges (2016-2018). Data source: SSA (2020); author's calculations.

However, this aggregated approval rate obscures the various paths judges take to meet that quota of 500 cases per year. Epstein et al. note that a judicial ruling classified as conservative may be motivated by ideology, but they theorize that it can also be motivated by legalism, strategic maneuvering, pragmatism, or even laziness (2013, p. 47). Likewise, the subset of judges who happened to adjudicate just over 500 disability appeals cases per year arrived there for a variety of reasons. Indeed, not all ALJs who are clustered around 500 dispositions are time constrained or even incentivized by the quota.

Consider this taxonomy: First, there is a subset of judges adjudicating just fewer or just over 500 cases per year who were not affected at all by the quota. In the absence of the quota, the distribution of the graph suggests there is a subset of judges who would have naturally adjudicated between 500-520 cases per year. If we reconstructed the above Figure 2.2 to show the distribution of judicial adjudications without a 500-case quota, there would still be a number of judges who would adjudicate 500-520 cases each year.

Second, it is possible that there is a subset of judges whose actions were affected by the quota, but they were negatively incentivized to decrease the number of cases adjudicated! In the absence of the quota, they would have adjudicated more than 520 cases.

Finally, there appears to be some judges who were in fact time constrained by the quota and incentivized to increase the number of cases they adjudicated to fulfill the quota of 500 cases per year. In the absence of the quota, they would have adjudicated fewer than 500 cases. The dataset I have allows me to construct a proxy metric to identify these judges who seem to be *time constrained* by the quota. Figure 2.3 disaggregates approval rates by those who are rushed and those who are not rushed in the last month of the fiscal year, as denoted by an increase in adjudicated cases by 30%.

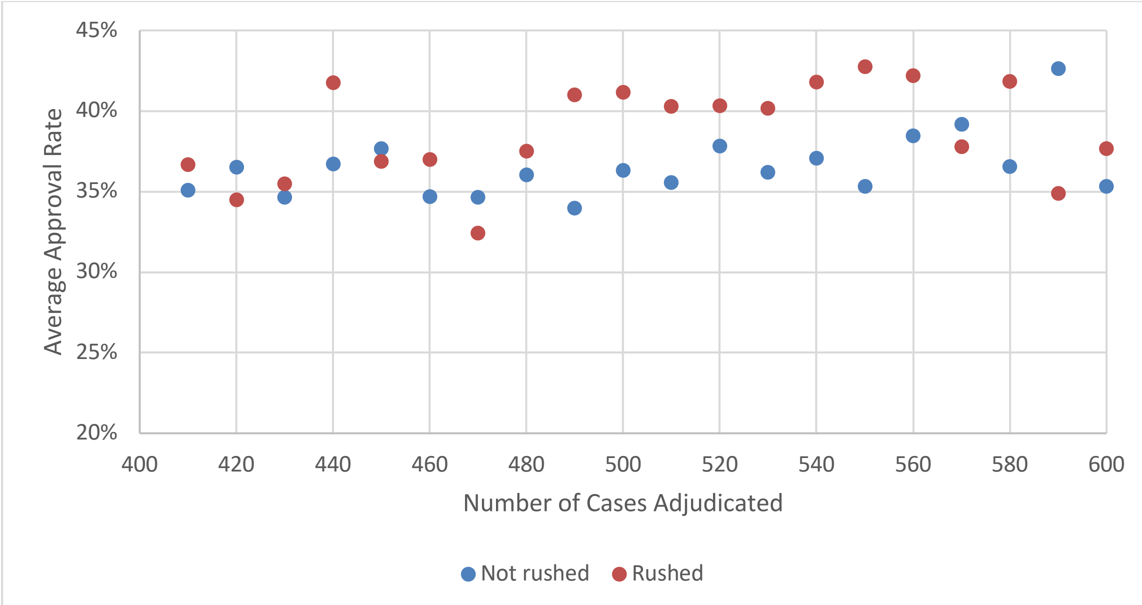


Figure 2.3. Administrative Law Judge Approval Rates Disaggregated by Rushed versus Not Rushed Judges (2016-2018). Data source: SSA (2020); author's calculations.

In a fractional response logistic regression model, these ALJs have a slightly higher approval rate than the rest of the ALJ population, and a slightly higher increase in approval rates in the last month of the fiscal year. Specifically, judges who are near to and appear to be constrained by the quota have a 2.1% higher annual approval rate and a 3.7% higher approval rate in September, the last month of the fiscal year. This translates to an increase of 9 approved cases per affected judge per year and roughly 2 approved cases in the last month of the fiscal year.

The rest of the chapter is organized as follows: Section 2 is a literature review, Section 3 presents the empirical strategy, Section 4 summarizes data used, Section 5 explains the empirical specification, Section 6 presents results, and Section 7 concludes.

2.2 Literature Review

A large body of literature suggests nonjudicial exogenous factors affect judicial decision making. For example, variations in outside air temperature (Heyes & Saboranian, 2019) led to a higher deportation rate amongst US immigration judges. Favorable parole decisions increased after a judge returned from a meal break (Danziger et al., 2011). Similarly, a judge's previous decisions appear to affect judicial decision making: Judges reviewing asylum court cases are "3.3 percentage points more likely to reject the current case if they approved the previous case" (Chen et al., 2016, p. 1182). Judges are people, too.

However, there is less literature on whether or how changes in judicial processes, such as changes in workload, impact judicial decision making. Esptein et al. find no correlation between judicial workload and rate of dismissals in district courts (2013, p. 229). However, they do find a statistically significant negative correlation between dissent rate and workload. Best and Tiede (2015) find that an increased federal district court judicial workload caused by an increase in vacancies led to harsher sentencing in criminal cases.

There are at least two reasons why the above findings related to judicial workload may not properly predict how workload changes, such as institution of a case adjudication quota, will affect disability appeals decisions. Federal district court judges face different incentives than administrative law judges when deciding a case. For most federal district court judges, any decision, positive or negative, is subject to an appeal. If a case is appealed and overturned, this could negatively affect a judge's reputation and may negatively affect workload because there is now a record that the judge decided a case incorrectly, which might affect the judge's reputation. In addition, decisions that are overturned are generally remanded to the district judge who decided the case originally, which leads to additional work (Epstein et al., 2013).

The incentives are different for an administrative law judge hearing disability benefit appeals.[15] Negative decisions are subject to further appeal, but affirmative ones are not. So there is no clear negative reputational effect if a judge approves an appealed case, whereas there is some unknown probability, x, that a negative decision could. Therefore, to the extent that administrative law judges are worried about having an opinion overturned on appeal like federal district court judges are (Choi et al., 2012), then approving an appeal eliminates this concern. This could incentivize administrative law judges to approve more cases to avoid negative reputation effects or to save time since the probability of reviewing an approval is lower, 0, than a rejection. Given the differences in incentives, it is not clear that effects of workload on judicial outcomes for federal district court judges are the same as for administrative law judges.

Nakosteen and Zimmer's (2014) study is one of the few to examine whether workload affects judicial decision making of SSA administrative law judges. Specifically, they test two distinct workload measures. First, they test whether the number of cases a

[15] Epstein et al. (2013) note that most of the research on judicial decision making and political affiliation focuses on Senate-confirmed Article III judges (p. 33).

judge adjudicates each year is correlated with approval rates and find no statistically significant correlation. Second, they test whether a judge's office-level backlog of cases affects approval rates, hypothesizing that a large backlog of cases at a judge's office may apply pressure for judges to rush through cases; once again, they do not find any correlation between office backlog of cases and approval rates.

However, Nakosteen and Zimmer (2014) do not test whether a quota that applies to a particular judge might influence judicial decision making. In fiscal year 2008, the SSA instituted a decision quota specifically designed to reduce office backlog data: ALJs need to decide 500-700 cases per year (SSA, 2012).[16] With the quota, SSA turned a general office case backlog into a personal case backlog. A judge's workload is no longer nebulously some portion of the backlog of the entire office, but rather it is specifically measured against the 500 cases he or she is advised to adjudicate each year.

Furthermore, not all judges interpreted this quota as an increase in workload. As Table 2.1 makes clear, 65% of judges in my sample did not meet the quota in fiscal years 2016-2018. This is not to say they were not affected by the quota. It is possible that they increased the number of cases they adjudicated even though they did not successfully meet the target of 500 cases. Second, some judges would have adjudicated just over 500 cases with or without this quota. Third, paradoxically, some judges could even be negatively incentivized by the quota. These judges may have decreased the number of

---

[16] An ALJ sued over this quota, claiming it violated judicial independence and forced them to approve more cases; the case lost in 2015 (Dirks, 2015).

cases adjudicated once they received formal policy that it was acceptable to adjudicate 500 cases per year. This would be consistent with other studies that find turning a norm into a price could have the opposite intended effect (Gneezy & Rustichini, 2000). In other words, although their actions could be impacted by the quota, they are not time constrained, or rushed, by the quota.

Finally, there are some judges who view the quota as an increase in workload. They are time constrained by the quota and pressured to meet it. The focus of this paper is on this subset of judges who are *near to* and *constrained by* the quota. I predict that this subset of judges will have a higher approval rate.

2.3 Empirical Strategy

Epstein et al. (2013, p. 48) create a judicial utility function for Senate-confirmed, Article III judges:

$$U = U(S(t_j), EXT(t_j, t_{nj}), L(t_l), W(t_j), Y(t_{nj}), Z)$$

Where a judge maximizes utility subject to time constraints: t_j is time working on the bench, t_l is leisure, and t_{nj} is nonjudicial activity that is adjacent to and correlated with her role as a judge, such a public appearances or journal writing: $t_j + t_{nj} + t_l = Total\ time\ in\ a\ day.$ S is job satisfaction, EXT is external satisfaction from being a judge, L is leisure, W is wage from being a judge, Y is income from nonjudicial activity, and Z a composite of all other factors.

Epstein et al. (2013) analyze how exogenous changes, such as increased workload, will affect the discrete elements of the utility function to predict how a judge will respond. Administrative law judges are Article II judges and although the utility function's broad elements still apply, the weights are different. For example, EXT and Y may still exist for Article II judges, but they play a much smaller role in their utility maximization.

With this framework, a quota will affect S, L, W, and Y, but it will affect different judges differently. For those judges adjudicating slightly fewer than 500 cases a year, this could lower S, L, and Y. They may feel compelled to hit the quota if the threat of firing is an option, affecting W. I could not get anyone to confirm or deny what effects quotas have on job security or pay, but it does not appear that judges are being fired; otherwise, we would see a lot of layoffs each year because well over 50% of judges are not achieving the quota.

However, while Figure 2.1 implies that some judges were affected by the quota, it is also clear from the figure that many judges did not meet the quota. Roughly 65% of judges fall below 500 case adjudications per year. Perhaps they were incentivized to increase from 430 cases per year to 440 cases per year. Data limitations prevent me from identifying this. It implies that the threat of being laid off for failing to meet the quota is quite low. The quota may also negatively affect those judges adjudicating well below 500 cases per year by inducing them to increase the number of cases adjudicated even if they do not come close to the quota, but this too is outside the scope of this paper because of data limitations.

For those judges who were adjudicating well over 500 cases a year, the introduction of a quota may incentivize them to reduce their workload. Reducing their workload increases S, Y, and L, while holding W unchanged. Now that the employer has signaled a minimum workload that is deemed satisfactory, some may reduce their workload since adjudicating more cases, expending more effort and time, does not change wage or employer approval. This is reminiscent of on the literature of extrinsic and intrinsic motivation.[17] For example, Gneezy and Rustichini (2000) found that transforming a volunteer task, i.e. asking for donations for a charity, into a paid job, actually reduced effort unless the extrinsic reward was high enough. So too with a quota. Prior to the quota, each judge determined the appropriate amount of effort to exert. Once the employer defined the appropriate level of effort, 500 cases, some judges who were well above the approved workload might be incentivized to reduce their effort, and therefore increase utility, without affecting W or EXT.

Although difficult to measure and outside the scope of this paper, introduction of a quota may also affect S more generally by creating an overly monitored, distrustful work environment. The ALJ lawsuit (Dirks, 2015) suggests that it affected at least some ALJs' morale given that they sued. This too is beyond the scope of this paper.

[17] For a review of the literature, see Deci et al. (1999).

2.4 Data

There was one main source of data for this chapter: the Social Security Administration's Administrative Law Judge (ALJ) Disposition Data for fiscal years 2016-2018 (SSA, 2020). Federal fiscal years begin October 1st and end September 30th. This dataset includes the judge's name, office, total dispositions at her current office and at any other offices, dismissals, decisions, awards, denials, and the number of awards that were fully or partially favorable by fiscal year and month. Cumulative results are reported monthly so I can derive number of decisions and approval rates by month. There are 184 offices, 10 regions, and approximately 1,600 ALJs deciding nearly 700,000 cases every year.

In the results reported below, I excluded individuals who had heard fewer than 300 cases in a year. Functionally, this means I excluded newly hired ALJs who started late in a given fiscal year and ALJs retiring at the start of a fiscal year. I also excluded judges who adjudicated fewer than 10 cases in September. For simplicity, I assigned each ALJ to one office per fiscal year and excluded ALJs' disposition data from other offices within a given year. I was left with 897 administrative law judges. The majority of these had 3 full years of data, but I included about 79 judges who had only 1 or 2 years of data. If I excluded these 79 from my analysis, results do not substantively change.Table 2.2 reports summary statistics of judge characteristics: 64% of judges are male and 36% are female; 46% are registered Democrats, 31% are registered Republicans, and 23% are Independent.

Table 2.2. Summary Statistics of Sample's Administrative Law Judges: Gender

Gender	Freq.	Percent	Cum.
female	330	36.79	36.79
male	567	63.21	100.00
Total	897	100.00	

I ensured that the decision to dismiss or hear a case was not different for judges near to the quota. First, I disaggregated the number of cases adjudicated by two different types of adjudication methods: dismissals or decisions. A judge can dismiss a case for a variety of reasons. For example, if an appellant misses a hearing date, the judge may dismiss the case.[18] On average, 20% of adjudicated cases are dismissals. Therefore, I predict a slightly positive increase in dismissals as adjudications increase. Figure 2.4 plots average number of dismissals by cases adjudicated. The number of dismissals does not appear to be affected by the quota.

[18] For the full list of reasons a case is dismissed, see HALLEX I-2-4-5 (SSA, 2014).

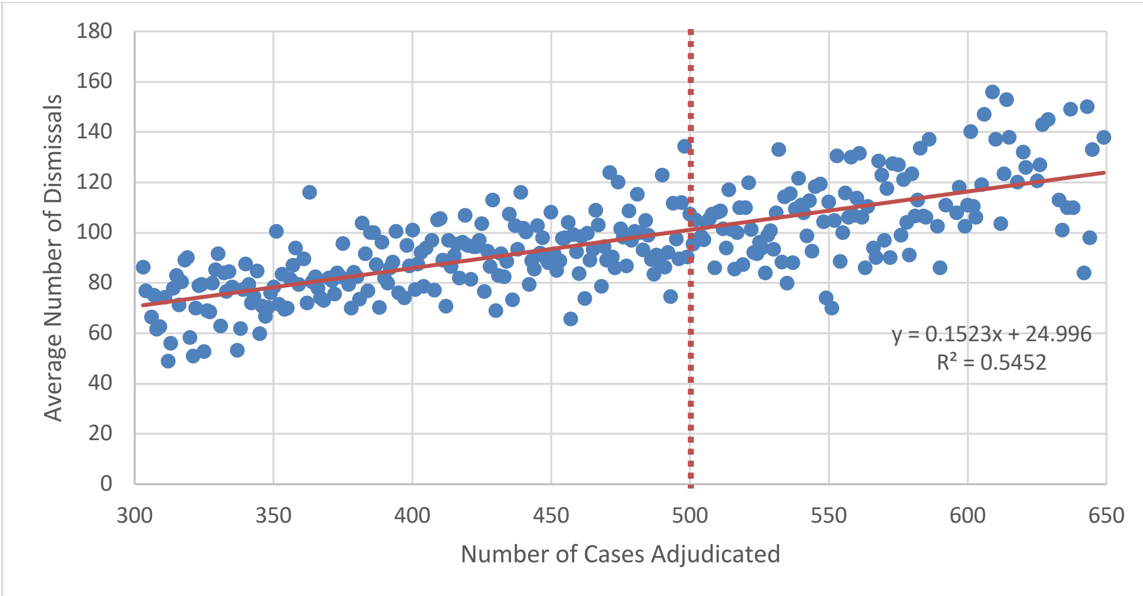


Figure 2.4. Number of Average Dismissals by Total Number of Administrative Law Judge Adjudicated Cases, 2017, author's calculation.

The union representing administrative law judges, the Association for Administrative Law Judges, sued SSA over the quota (*Association of Administrative Law Judges vs. Colvin*, 2001), claiming that a quota would affect how they decided cases. Specifically, they claimed that because it is more difficult to deny than accept an appeal—because denials must withstand scrutiny in the case of appealing the decision—judges will be nudged to accept appeals to hit the quota. Given that the Administrative Procedure Act (APA) of 1954 gave administrative law judges quasi-independence from their respective agencies to ensure quality control,[19] they argued that rule is therefore illegal.

Using a logistic regression, I test whether any personal characteristics predict whether the judge will make the 500-case adjudication goal (Table 2.3, Table 2.4).

$$\mathrm{Ln}[\frac{P\_500goal}{1-500goal}] = \alpha + b_1 approvalrate + b_2 gender + b_3 (registered\ pol.party) + b_4 age + b_5 fiscalyear + b_6 office + b_7(fiscalyear * office) + \varepsilon \quad [2]$$

Table 2.3. Results from Logistic Regression of 500-Case Goal on Administrative Law Judge Judicial Characteristics

	(1) Make 500 quota
Approve-rate	2.093***
	(3.85)
Male	0.652***
	(5.15)
Republican	0.0352
	(0.25)
Independent	0.200
	(1.25)
age	-0.0163*
	(-2.38)
N	1836

[19] As of July 2018, they are no longer considered independent from the agency. Another court case is described in Christiansen (2018).

t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 2.4. Odds Ratio

	(1)
	greatert~500
approve-rate	8.113***
	(3.85)
Male	1.920***
	(5.15)
Republican	1.036
	(0.25)
Independent	1.222
	(1.25)
age	0.984*
	(-2.38)
N	1836

Exponentiated coefficients; t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Higher approval rates statistically significantly predict whether a judge makes the quota. Similarly, male judges are nearly twice as likely to achieve the quota and decide at least 500 cases than female judges. Political affiliation is not correlated with meeting the quota.

However, if I artificially change the quota to 450 or 550, the same results hold for gender; male judges are more likely than female judges to make these fabricated goals as well, suggesting that, on average, male judges decide more cases than female judges. Interestingly, the odds ratio of approval rates increases from roughly 8 to 28 when I create an artificial quota of 550.

I then further disaggregated my analysis to focus on those who appear to be not only near the quota, but also time constrained and incentivized by the quota. To do this, I created a proxy for judges who are near to and time constrained by the quota. First, I labeled any judge adjudicating 500-520 cases a year as a judge who is near the quota. To code those who are time constrained by the quota, I identified the subset of judges who adjudicated a large number of cases at the end of the fiscal year. In order to adjudicate 500 cases a year, an ALJ must adjudicate on average **41** cases each month. I defined rushed quota as an ALJ who adjudicated between 500 and 520 cases in a year and who decided at least **54 cases** in September, a 30% increase in average monthly adjudications. I experimented with higher and lower cut points for both the total number of cases adjudicated and the September number of cases adjudicated and it did not substantively change results. Given this classification, 157, or just over 6% of my sample of administrative law judges, were near to and constrained by the quota.

Figure 2.5 displays the number of judges who are rushed by the total number of cases adjudicated. Once again, we see a spike in the number of judges who are rushed for judges who adjudicated between 500 and 520 cases. As Table 2.5 reports, 19% of judges in my sample adjudicate between 500 and 520 cases; 37% of those judges are rushed. It should also be noted that there is a high percentage of rushed judges who adjudicate 480-500 cases as well. This makes sense: Some judges may have rushed to meet the quota but failed to adjudicate 500 cases because of unforeseen circumstances. In this chapter's results section and Appendix B, I explore different proxies for rushed judges to see if my results change.

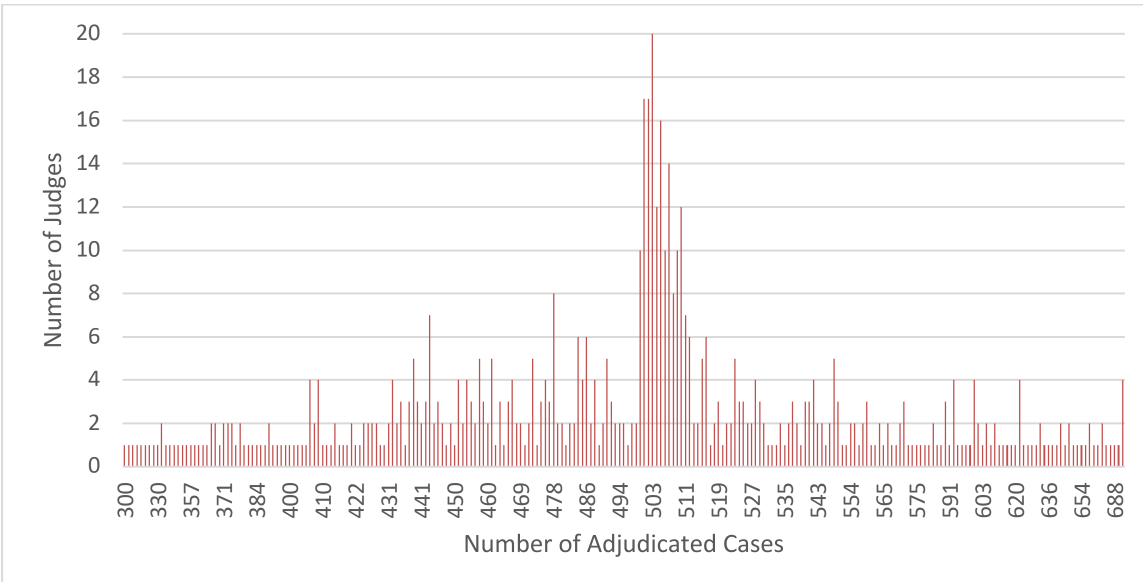


Figure 2.5. Number of Rushed Administrative Law Judges by Number of Cases Adjudicated, 2016-2018, author's calculation.

Table 2.5. Number of Rushed Administrative Law Judges by Cases Adjudicated, 2016-2018

Number of Adjudicated Cases	Number of Judges	Percentage of Judges	Number of Rushed Judges	Percentage of Rushed Judges
300-400	421	17.7%	48	7.9%
400-410	93	3.9%	16	2.6%
411-420	83	3.5%	9	1.5%
421-430	67	2.8%	12	2.0%
431-440	93	3.9%	23	3.8%
441-450	91	3.8%	23	3.8%
451-460	114	4.8%	25	4.1%
461-470	100	4.2%	22	3.6%
471-480	99	4.2%	30	4.9%
481-490	69	2.9%	30	4.9%
491-500	78	3.3%	29	4.7%
501-510	**317**	**13.3%**	**136**	**22.3%**
511-520	**143**	**6.0%**	**35**	**5.7%**
521-530	132	5.6%	28	4.6%
531-540	73	3.1%	17	2.8%
541-550	75	3.2%	22	3.6%
551-560	68	2.9%	12	2.0%
561-570	45	1.9%	11	1.8%
571-580	33	1.4%	8	1.3%
581-590	35	1.5%	8	1.3%
591-600	26	1.1%	13	2.1%
greater than 600	122	5.1%	54	8.8%

2.5 Specification

$$\text{Ln}[\frac{P\_APP}{1-APP}] = \alpha + b_1 RushedQuota + b_2 gender + b_3\ (registered\ political\ party) + b_4 age + b_5 (number\ of cases) + b_6 fiscalyear + b_7 office + b_8 (fiscalyear * office) + \varepsilon$$

[3]

2.6 Results

Equation 3's results are displayed in Table 2.6 and

Table 2.7, controlling for office and fiscal year. Table 2.6 shows a comparison between everyone in the dataset and just these few folks rushing to deadline. Being near to and ostensibly rushed by the quota increases a judge's approval rate by 2.1% for the entire year. Table 2.8 shows the margins.

Table 2.6. Fractional Response Logistic Regression of Administrative Law Judge Approval Rate on Rushed

	(1)	(2)	(3)	(4)	(5)
Rush500-510	0.0702 (1.76)				
Rush500-520		0.0918* (2.51)			
rush500-530			0.0949** (2.75)		
rush500-550				0.111*** (3.36)	
rush480-520					0.0874** (2.65)
Male	0.101*** (4.84)	0.102*** (4.88)	0.102*** (4.89)	0.102*** (4.88)	0.102*** (4.90)
Republican	-0.167*** (-7.01)	-0.167*** (-7.02)	-0.167*** (-7.00)	-0.168*** (-7.07)	-0.167*** (-7.01)
Independent	-0.204*** (-8.09)	-0.205*** (-8.14)	-0.206*** (-8.18)	-0.207*** (-8.21)	-0.206*** (-8.16)
age	0.0112*** (10.15)	0.0112*** (10.15)	0.0113*** (10.17)	0.0113*** (10.19)	0.0112*** (10.14)
Cases	0.000468*** (3.57)	0.000456*** (3.52)	0.000449*** (3.49)	0.000430*** (3.41)	0.000453*** (3.51)
N	2377	2377	2377	2377	2377

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 2.7 Odds Ratio of Fractional Response Logistic Regression for Rushed 500-520

	(1)
rushed500-520	1.096*
	(2.51)
Male	1.107***
	(4.88)
Republican	0.846***
	(-7.02)
Independent	0.815***
	(-8.14)
age	1.011***
	(10.15)
cases	1.000***
	(3.52)
N	2377

Exponentiated coefficients; t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 2.8. Margins

	dy/dx (Delta-method)	Std. Err.	z	P>|z|
Rushed500-520	.0208325	.00829	2.51	0.012
male	.0231596	.0047445	4.88	0.000
Republican	-.0381182	.0053975	-7.06	0.000
Independent	-.0465494	.0056536	-8.23	0.000
age	.0025531	.0002513	10.16	0.000
cases	.0001036	.0000294	3.52	0.000

Note: dy/dx for factor levels is the discrete change from the base level.

I performed the same analysis as above, but focused on September, the last month of the fiscal year. The dependent, or response variable, is now a judge's September approval rate, rather than his annual approval rate. Table 2.9 reports coefficients, Table 2.10 reports the odds ratio for (1), and

Table 2.11 reports margins for (1). Results are similar, but stronger. Approval rates for those constrained by and near to the quota have approval rates 3.7 percentage points higher than all other judges in the sample.

Table 2.9 Fractional Response Logistic Regression of Administrative Law Judge September Approval Rate for Rushed Judges

	(1)	(2)	(3)	(4)
rush500~510	0.161**			
	(2.96)			
rush500~520		0.171***		
		(3.49)		
rush500~530			0.0949**	
			(2.75)	
rush500~550				0.153***
				(3.49)
Male	0.157***	0.158***	0.102***	0.158***
	(5.57)	(5.63)	(4.89)	(5.63)
Republican	-0.179***	-0.179***	-0.167***	-0.180***
	(-5.56)	(-5.55)	(-7.00)	(-5.58)
Independent	-0.231***	-0.233***	-0.206***	-0.233***
	(-6.73)	(-6.77)	(-8.18)	(-6.77)
age	0.00971***	0.00969***	0.0113***	0.00972***
	(6.62)	(6.60)	(10.17)	(6.63)
Cases	0.000316	0.000300	0.000449***	0.000276
	(1.79)	(1.72)	(3.49)	(1.61)
N	2377	2377	2377	2377

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Table 2.10. Odds Ratio for September Approval Rate, Rushed, Deciding 500-520 Cases

	(1)
rushed500~520	1.186***
	(3.49)
Male	1.172***
	(5.63)
Republican	0.836***
	(-5.55)
Independent	0.792***
	(-6.77)
age	1.010***
	(6.60)
cases	1.000
	(1.72)
N	2377

Exponentiated coefficients; t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 2.11 Margins for September Analysis, Rush 500-520 Cases

	dy/dx	Delta-method Std. Err.	z	P>|z|
rush500to520	.0376075	.0107522	3.50	0.000
male	.0348941	.0061818	5.64	0.000
Republican	-.039712	.0071051	-5.59	0.000
Independent	-.051165	.007465	-6.85	0.000
age	.0021363	.0003233	6.61	0.000
cases	.0000662	.0000384	1.72	0.085

Note: dy/dx for factor levels is the discrete change from the base level.

2.6.1 Robustness checks

Results are similar if I perform the same analysis with a linear regression (Table 2.12).

Table 2.12. Linear Regression Rather Than Fractional Response Method for Those Who Rushed and Adjudicated 500-520 Cases

	(1)
rushed50~520	0.0211*
	(2.22)
gender	0.0234***
	(4.45)
Republican	-0.0382***
	(-6.32)
Independent	-0.0468***
	(-7.42)
age	0.00258***
	(9.11)
Cases	0.000108***
	(3.31)
N	2377

t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

I also tested the effect of being rushed in September, independent of how many cases the ALJ decided, and found that being rushed in September is statistically significantly and positively correlated with an increase in an ALJs' approval rate in September, but not an increase in his approval rate for the entire year. I also tested whether being rushed near to and constrained by the quota affected individuals differently based on gender or political affiliation and did not find any statistically significant correlation. In Appendix B, I include another variation to test robustness. The results hold and remain statistically significant.

2.7 Conclusion

In their lawsuit against the SSA Commissioner, the AALJ asserted that an adjudication goal would affect judicial decision making (*Association of Administrative Law Judges vs. Colvin*, 2001). The above analysis suggests that, for a small subset of the ALJ population, it does. The magnitude of the effect is small, 2.1% increase in approval rates throughout the year and 3.7% increase in the final month of the fiscal year, for a small subset of judges, 6%. This translates to roughly 11 cases per year per affected judge. Given that there are roughly 1,600 SSA administrative law judges, this translates to an increase of about 1,000 cases approved each year. Of these approvals, roughly 150 will occur in the final month of the fiscal year.

There is a tension between timeliness and accuracy. My proxy for identifying those who are time constrained and incentivized by the quota may be imperfect given that I never received formal confirmation of the quota's details. Future research should experiment with better measures to identify those who are affected by the quota.

3. WHAT DO SMALL PROCUREMENT FIRMS DO?

3.1 Introduction

Recent research has illuminated several key facts about business dynamism. First, new firms rather than small firms are the engines of economic growth (Haltiwanger et al., 2010). Most new firms fail, and for the relatively few that do succeed, most remain small in size and participate in industries amenable to small-scale operations (Hurst & Pugsley, 2012).

Unfortunately, neither the American imagination nor its regulation has caught up with these realities. According to a 2011 Gallup Poll (Newport, 2011), Americans trust small businesses to create jobs over economists, the president, the federal reserve, and large businesses. In Gallup's annual institutional confidence poll (*Confidence in Institutions*, n.d.), small businesses, along with the military, have been perceived as one of the most trusted of American institutions since the poll's inception in 1997. In fact, according to the poll, in 2017, Americans' positive perception of small businesses was at an all-time high of 70%. American politicians of all stripes continue to laud the small-business owner as the backbone of the American economy and back up these words with generous policies favoring small businesses even though evidence suggests such subsidies are likely regressive (Hurst & Pugsley, 2012, 2015) and ineffective (Haltiwanger et al., 2014).

One such subsidy afforded small businesses today is preference in the federal procurement market[20]. Given the finding that most small businesses tend to stay small (Hurst & Pugsley, 2012), it is questionable whether this preference is helpful to the procurement market. More importantly, given that the federal procurement market's structure and incentives are distinct from the regular economy, it is unclear whether firms providing goods and services to the federal government—be they small or large, young or old—have the same business dynamism as those which participate in the regular economy. Since the federal government has explicit programs to favor small businesses in government contracting, we should expect to see entrepreneurs enter this space to provide goods and services to the federal government. I offer the first-ever set of stylized facts that small businesses in the procurement economy are indeed far different in size, industry, and life cycle from small businesses across the rest of the U.S. economy. This is a sign that these small businesses may in large part be the creation of government procurement rules rather than a product of the market process.

Relatively little is known about firm dynamism within the government procurement economy despite the fact that it makes up a nontrivial portion of the U.S.

---

[20] The preferences are numerous and slice the SB market for Federal contracts into subdivision – each striving to win the targeted 23% of annual Federal contract obligations. See FAR Part 19.

economy and has attributes that may make it different from the normal economy. The federal government spends roughly $550 billion annually on procurement; state and local governments spend roughly $1 trillion annually (Farmer, 2019). Together, this is roughly 7-8% of the U.S. economy.

Moreover, firms in the federal procurement economy appear to organize themselves differently from the regular economy. Consider this fact: Over 50% of all federal government procurement dollars are awarded to fewer than 100 firms.[21] This market power appears to mimic the market power of top firms in the regular economy, but in any given year, each of these top firms in the procurement economy is highly diversified and provides goods and services spanning 20-30% of all possible industries. Such evidence suggests that firms may organize differently and specifically to serve this market.

Finally, there is additional government procurement-specific regulation that could cause this market to behave very differently from the rest of the economy. Firms wishing to work with the federal government must comply with all standard U.S. regulations in addition to this procurement regulation.

The focus of this chapter is to begin investigating procurement firm demographics and dynamism by exploring whether small business demographics and dynamism in the U.S. economy hold for businesses which participate in the federal

---

[21] Obtained from Federal Procurement Data System – Next Generation *Top 100 Contractors Reports* for FY2006-2018 (FPDS-NG, n.d.).

procurement economy. Given the structure and added regulation, what industries do small firms participate in and what predicts their success? To do this, I first look at concentration of small firms by industry in the federal procurement economy, similar to what Hurst and Pugsley (2012) do for the overall U.S. economy. This allows me to illustrate the differences in the nature of small businesses in these two worlds. I then identify firm and contract-specific attributes that are correlated with small firm success in the procurement economy.

I find that small procurement firm demographics and dynamism share similarities to small U.S. firms in the regular economy but they deviate in some key ways. Small procurement firms are similar to all small U.S. firms in that they are concentrated in a narrow range of industries: 65% of all small procurement firms participate in just 40 industries, out of a possible 312. Additionally, although the specific industries in which small procurement firms cluster are not identical to small U.S. firm industry clustering, 17 of the top 40 industries in the procurement economy are the same as the top 40 industries for the regular economy.

However, there are some notable differences in small procurement firm industrial composition. Specifically, small procurement firms participate in manufacturing industries at much higher rates than do all small U.S. firms: Nearly 25% of the top 40 industries that small procurement firms participate in are manufacturing related. This is in stark contrast to all small U.S. firms: Not a single manufacturing industry is listed in the top 40 industries for all small firms.

Additionally, the procurement industries in which small procurement firms are heavily concentrated are not dominated by these small firms. There is little difference in an industry's share of small business, as measured by share of the industry's contract dollars, between the industries with the highest concentration of small businesses and those with the lowest concentration of small businesses. This too is a departure from research on small firm industry participation (Hurst & Pugsley, 2012) which found that small businesses are concentrated in service industries such as restaurants, law firms, and doctor's offices, **and** that these industries have a higher share of small firm employment than industries with low small firm concentration, suggesting that it is the structure of the industry rather than the size of the industry that is driving small firm concentration in those industries.

Finally, it is quite common for small firms in the procurement economy to span several industries; crucially, this participation in several different industries, as measured by 4-digit North American Industry Classification System (NAICS) codes, is strongly correlated with small firm success in the procurement economy. Specifically, for small firms registering with the federal government in fiscal year 2013, winning a contract in an additional 4-digit NAICS code increases contract dollars obligated by 48%. While there are several possible explanations for this phenomenon, I argue that winning contracts spanning several NAICS codes is a proxy for insider procurement knowledge, that is, it is a proxy for specialization in procurement regulation, processes, and networks, rather than authentic industry diversification. I do this in part by comparing firm owner's self-reported NAICS specialization in the government's business repository with another online business repository that is not used in federal procurement processes. Successful procurement firms are more likely to self-report many 4-digit NAICS codes in federal documents but, according to company filings interpreted by private sector business databases, they rarely present more than one 6-digit NAICS code.

The chapter is organized as follows: In Section 2, I review the literature on small and new business dynamism in the regular economy. In Section 3, I provide high-level industrial composition data of small U.S. procurement firms and contrast it with similar data for all small U.S. businesses (Hurst & Pugsley, 2012). In Section 4, I present preliminary analysis on predictors of small firm success in the procurement economy. In Section 5, I conclude by summarizing findings and identifying areas for future research.

3.2 Literature Review and Data Sources

We know very little about the industrial composition of small firms in the procurement economy. This is in part because the federal procurement economy has a different definition of small business than the private sector. Hurst and Pugsley (2012) define small businesses as having fewer than 20 employees, and I use their definition. In the procurement economy, the definition of small business varies by industry and tends to be more expansive than the definition of small business in the economic literature.[22]

[22] For a larger discussion of the definition of small business in the federal procurement economy, please see Appendix C.

Therefore, most analysis of small businesses in the procurement economy is not a perfect apples-to-apples comparison. However, we know quite a lot about small firm industrial composition in the regular economy. Hurst and Pugsley (2012) find that for the overall U.S. economy, 60% of small firms are clustered in 40 4-digit NAICS codes. Given that there are roughly 300 4-digit NAICS industry categories,[23] this means that 60% of small firms are clustered in just 14% of industries and 80% of small business are clustered in just 30% of industries.

Furthermore, according to Hurst and Pugsley (2012), all of these top 40 small business industries are services and most of these industries fall into one of the following more general categories: restaurants, gas stations, skilled craftspeople, skilled professionals, small shopkeepers, and doctor's offices. Moreover, these 40 industries in which small firms tend to cluster also have the highest concentration of employment in small firms, suggesting that these industries are amenable to small-scale operations. Finally, firms in these top 40 industries do not seem to grow in employment over time, providing more evidence that on average these are industries where strongly diminishing returns to scale are the norm.

Should we expect to see the same type of concentration in the procurement economy? Many of these top 40 industries listed above, such as restaurants and gas

[23] The number of 4-digit NAICS codes changes gradually over time as the economy changes. For example, when Hurst and Pugsley (2012) analyzed small firms in 2003, there were 294 4-digit NAICS codes. When I performed my analysis, there were 312 NAICS codes. As of 2017, there were 311 4-digit NAICS codes (U.S. Census Bureau, n.d.), author's calculation.

stations, do not play a large role in the procurement economy, in which all transactions are business to "business," or rather business to government, rather than business-to-consumer (B2C). Does this change the structure of industries that are considered amenable to small-scale operations to being large scale?

Additionally, this understanding of the industries in which small businesses seem to cluster does not appear to be a factor when agencies are considering whether to contract with a small business. Contracting officers are required to do market research to see if small firms can provide a good or service, but the market research appears to be focused on historical success in the procurement economy, rather than in the regular economy. For example, a former U.S. Small Business Administration online course[24] for federal contracting officers identifies several resources for determining whether a contract can be set aside for small businesses, and almost all of those resources are focused on historical success in the procurement economy. This chapter will catalog the industrial composition of small firms in the procurement economy and compare and contrast that industrial composition with the regular economy.

Second, what type of firms will succeed in the procurement economy? We know that it is a small proportion of young firms, rather than small firms, that drive growth and dynamism in the regular economy (Haltiwanger et al., 2010). Hurst and Pugsley

---

[24] Market Research Guide for Contracting Officers, https://www.sba.gov/sites/default/files/files/mkt_transcript.pdf (U.S. Small Business Administration, 2013).

(2012) reinforce this point with qualitative data. Using surveys of small business owners, they find that the majority of small firms want to stay small: Most small business owners express little interest in growing large or entrepreneurship. Finally, Haltiwanger et al. (2016) identify high-growth industries for small firms.

Should we expect to see the same trends in the procurement economy? On the one hand, the government, when deciding to buy rather than make, is trying to leverage private sector expertise, innovation, and cost savings. This implies the government should want a firm with a proven track record in a particular commercial sector. It seems plausible that small, mature firms, rather than new and untested firms would be more attractive to a government agency. However, perhaps these small, mature firms are seeking government contracts for the wrong reasons. It could be they are turning to government contracts because they are not doing well with their current customer base and are seeking alternative revenue sources. The quality of mature, small firms who enter the procurement market may be lower. This chapter will test whether new firms rather than experienced small firms are more successful in the procurement economy.

However, there is an implicit assumption in the literature and the above analysis that small businesses participate in only one industry, as measured by 4-digit NAICS codes. But this industry composition in the procurement economy is not as clear cut as it is in the Hurst and Pugsley (2012) analysis. Firms in the procurement industry can and often do occupy several 4-digit NAICS codes.

There is historical precedence of firm diversification in government procurement. After WWII, military contract spending slowed and the military industrial complex had to adapt to remain viable in nonwar periods. The firms that were overly reliant on military contracts either had to diversify their customer base or the products they sold. Boeing chose to diversify its customer base and increase commercial sales while General Dynamics (originally named Electric Boat) chose to diversify its government offerings. By 1955, “97 percent of its [General Dynamics] sales went to the government, but its participation in five product lines for two government markets represented the greatest degree of diversification up to that time” (Nagle, 1999, p. 477).

In 2017, the top 10 firms in the procurement economy had federal contracts that spanned an average of 62 4-digit NAICS codes out of the total 312 NAICS codes, or roughly 20% of all possible industries. General Dynamics spanned 98 4-digit NAICS codes, or nearly 30% of all industries in FY2017. A simple scatterplot of log contract dollars to number of contracts with a distinct NAICS code is shown in Figure 3.1.

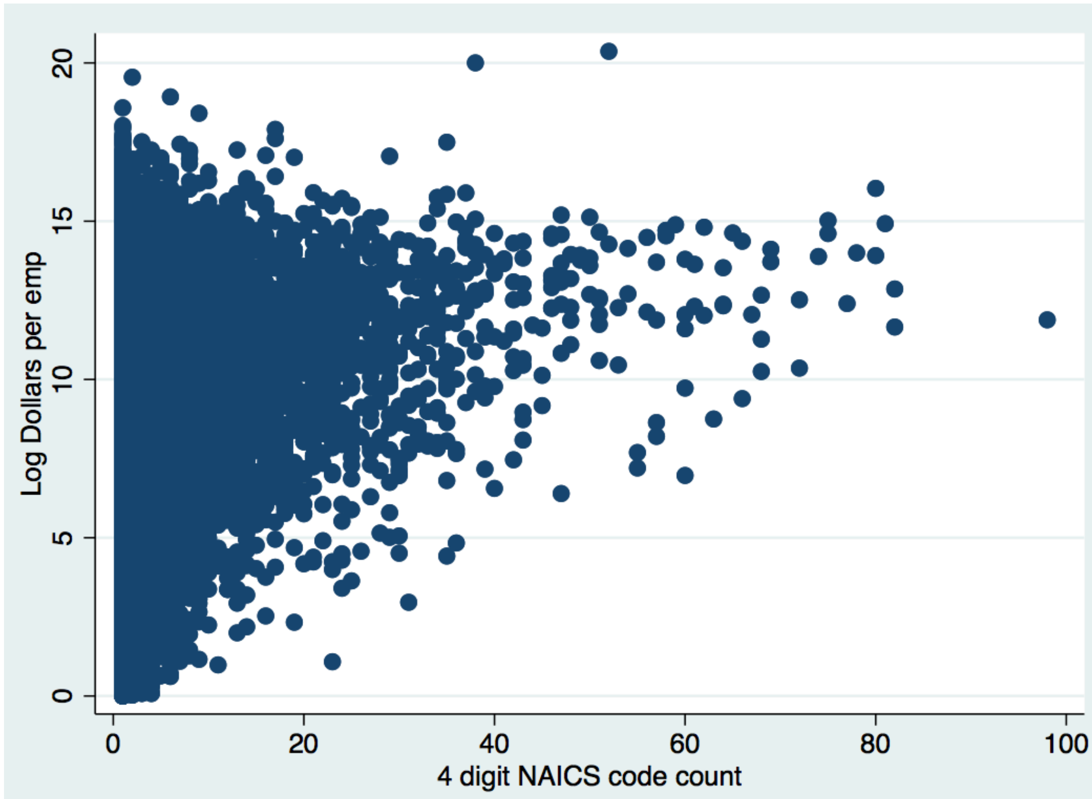


Figure 3.1. Scatter Plot of Log Contract Dollars and Number of NAICS Codes in FY17 for All Firms in the Procurement Economy, author's calculation.

Will successful small firms participating in the procurement economy behave similarly? Winning contracts in several different NAICS codes as a small firm could mean one of several things. It could signal authentic diversification, coding error, or a deeper understanding of the procurement economy.

One possibility is that this is actually how small businesses grow both inside and outside of the procurement economy, and we just do not have access to this granular level of data in the private sector. Then, if that is the case, using the federal procurement database to show how businesses grow could be helpful to better understanding all small business dynamism.

This type of growth is doubtful given the current literature on business diversification. It has been documented that in simple regression models, there is a negative relationship between business diversification and firm success (Martin & Sayrak, 2003). That relationship seems, however, to be due, in part, to endogeneity issues. That is, firms may be choosing to diversify because of negative performance (Shyu & Chen, 2009). After controlling for endogeneity, results are still mixed. For example, Shyu and Chen (2009) find that diversification has a slight positive effect on performance for *mature* firms who pursue related diversification (i.e., expanding in related products or services), rather than unrelated diversification (i.e., expanding in unrelated products or services). Young firms do not appear to gain a premium from diversifying.

A second, more likely possibility is that data quality within the federal procurement system is questionable. For example, contracting officers who select the 6-digit NAICS code for a specific contract requirement receive little training on selecting NAICS codes and therefore do not always select the right NAICS code for a contract. Another related possible problem is that firm owners are unsure of what NAICS code they belong to. Or, the coding error could occur when a firm is entered into the contract award database. One or more of these weaknesses with the procurement process seems entirely possible.

There is evidence to suggest that NAICS code selection is not always accurate. Naval Postgraduate School students sampled 276 Air Force contracts from 7 different 6-digit NAICS codes and found that contracting officers selected the correct 6-digit NAICS code 68% of the time (Miller & Ellis, 2016). To be clear, Miller and Elliis did not identify error rates at the 4-digit NAICS code level, but the error rate is almost certainly lower, if not substantially lower, than at the 6-digit level. I will test the validity of the claim that this is simply a coding error by comparing the NAICS codes associated with contracts won by small firms versus those NAICS codes that small firms claim represent their areas of expertise.

Finally, there is additional government procurement-specific regulation that could cause this market to behave very differently from the rest of the economy. Firms wishing to work with the federal government must comply with all standard U.S. regulations in addition to this procurement regulation. The *Federal Acquisition Regulation* (FAR), the set of regulations governing federal procurement, is well over 1,800 pages and continues to grow. Canada's equivalent, the Public Works and Government Services Canada's (PWGSC) *Standard Acquisition Clauses and Conditions (SACC) Supply Manual* (PWGSC, n.d.), is roughly 100 pages. The difference in size and complexity between these documents is somewhat shocking.

This comparison excludes the 70 agency-specific US FAR supplements that, for the largest agencies, rival the size of the original FAR.[25] The Department of Defense (DOD) is an extreme example: It not only has an agency-specific supplement called the Defense FAR Supplement (DFARS), but every single military department also has its own *supplement to this supplement*![26] I wish I were kidding, but I am not. For example, an Air Force contracting officer is governed by the FAR, the DFARS, and the Air Force FAR Supplement (AFFARS), and therefore must cope with over 4,000 pages of acquisition regulation.

Wilson declares the DFARS "of immense length, excruciating detail, and soporific prose" (2989, p. 241). According to a survey, 70% of business owners surveyed cited onerous paperwork and processes as the reason they would not consider DOD contracts (Lamm, 1988, p. 48). Furthermore, most cited paperwork and process in the pre-award stage as the reason they did not want to do business with DOD. In many ways, the federal procurement market provides a window into how an economy might behave if there were more regulation.

This chapter will test whether indicators correlated with firm success, such as firm age or specialization in high-growth industries, predict firm success in the procurement economy; I will also test whether indicators correlated with large firm success in the procurement economy, such as industry diversification, predict small firm

---

[25] For a list of agency-specific procurement regulation see eCFR (n.d.).
[26] There are regulation supplements at the military unit as well.

success in the procurement economy. Finally, I will provide preliminary evidence that small firm success may be tied to specialization in procurement regulation, processes, and networks rather than a particular industry.

3.3 Data Sources

Because of government transparency regulations, there is a detailed level of spending data in the federal government procurement economy. I use data from the Federal Procurement Data System – Next Generation (FPDS-NG, n.d.). This database provides a catalog of all federal spending at the contract-action level[27]. I aggregate this data from USAspending.gov (2018) to compile a list of firms that do business with the federal government, the amount of contract dollars awarded, the number of employees they report, and the NAICS code classification of the contracts they win and execute.

The biggest challenge with this database is that all data is reported by data universal numbering system (DUNS) numbers rather than establishment or firm. In Appendix C, I document how I tried to approximate a firm given the data limitations. Throughout this chapter, I refer to my unit of analysis as firm, but this is a proxy for my attempt to get at firm-level data.

I also employed the federal government's System for Award Management (SAM) database, the federal government's repository of all firms that have registered to bid on

---

[27] Contract transaction is defined as a transaction that obligates funds. See FAR Part 2 for a definition.

federal contracts and/or grants. SAM provides a variety of data points about the firm, but the one I use for this analysis is the firm's selection of NAICS code specialization. When registering in SAM, firms must specify at least one 6-digit NACIS code to identify their area of specialization, but can elect as many as they wish.

Finally, I used Mergent Intellect, a repository of all businesses, to augment or verify data. Mergent Intellect identifies firms by DUNS number, the main business identifier in federal procurement systems. I confirmed firm employment and augmented my dataset with firm age and their reporting of each business's primary and secondary industry by 6-digit NAICS code based on business filings with the Securities and Exchange Commission (SEC).

I link businesses between these three databases by their DUNS number.

3.3.1 North American Industry Classification System (NAICS) Codes

NAICS is one of many classification systems to identify businesses by industry (U.S. Census Bureau, 2020). There are over 1,000 6-digit and 313 4-digit NAICS codes. To give an example of what is included in one 4-digit NAICS code, I summarize services covered by NAICS code 5617: *services to buildings and dwellings*. This 4-digit NAICS code includes all janitorial services, all extermination and pest control services, all landscaping services, all carpet and upholstery cleaning services, and all other services to buildings and dwellings. The catchall phrase "other services to buildings and dwellings" includes but is not limited to services like swimming pool repair, drain and gutter cleaning, building exterior cleaning services, duct cleaning, and chimney cleaning.

3.4 Industrial Composition of Small Businesses in the Procurement Economy

Although there is a lot of literature on U.S. small business demographics and dynamism, very little is known about small businesses that participate in the federal procurement economy. I follow Hurst and Pugsley's (2012) methodology to provide a clearer picture of small procurement firm demographics.

In 2013, there were 5.8 million firms operating in the United States. Roughly 90% or 5.2 million of those firms had fewer than 20 employees (U.S. Small Business Administration Office of Advocacy, 2016).[28] Very few of these 5.8 million firms participate in the federal procurement economy in any given year. Just 2% of U.S. firms[29] have at least one contract with the federal government in any given year, but it helps to disaggregate further into small and large firms. Small business participation in the federal government procurement economy is relatively small. In any given year, roughly 1%[30] of all small firms participate in the federal procurement economy.

If we confine our analysis to small businesses' concentration within the federal procurement ecosystem, small procurement firms play a somewhat larger role. In Fiscal Year 2013, there were over 136,000 procurement firms providing goods and services to the federal government and nearly 60% of them were small (80,377, 59%; compared to

---

[28] Similar results are in the 2012 Statistics of U.S. Businesses (SUSB) annual data survey (U.S. Census Bureau, 2015); author's calculation.
[29] Author's calculation.
[30] Author's calculation.

55,665 large business, 6,900 zero-employee firms).[31] Although small procurement firms comprise well over half of all procurement firms, they earn only about 5% of all procurement dollars.[32]

Figure 3.2 shows a cumulative distribution function by firm size in the procurement economy, summarizes the results above, and provides additional information on revenue and employee count. The most striking aspect of this graph is that most procurement dollars are concentrated in very large firms having at least 2,500 employees. This mimics Hurst and Pugsley's (2012) cumulative distribution for the entire economy, although small firms account for a much larger share of the U.S. economy than they do in the federal procurement economy.

---

[31] USAspending.gov (2018), author's calculation.

[32] I have excluded nonemployer firms from these calculations. Small businesses are 56% of all procurement firms if nonemployer firms are included (6,900 nonemployer firms). Also please note again, I am using Hurst & Pugsley's (2012) definition of small business in these calculations, not the federal government's definition of small business.

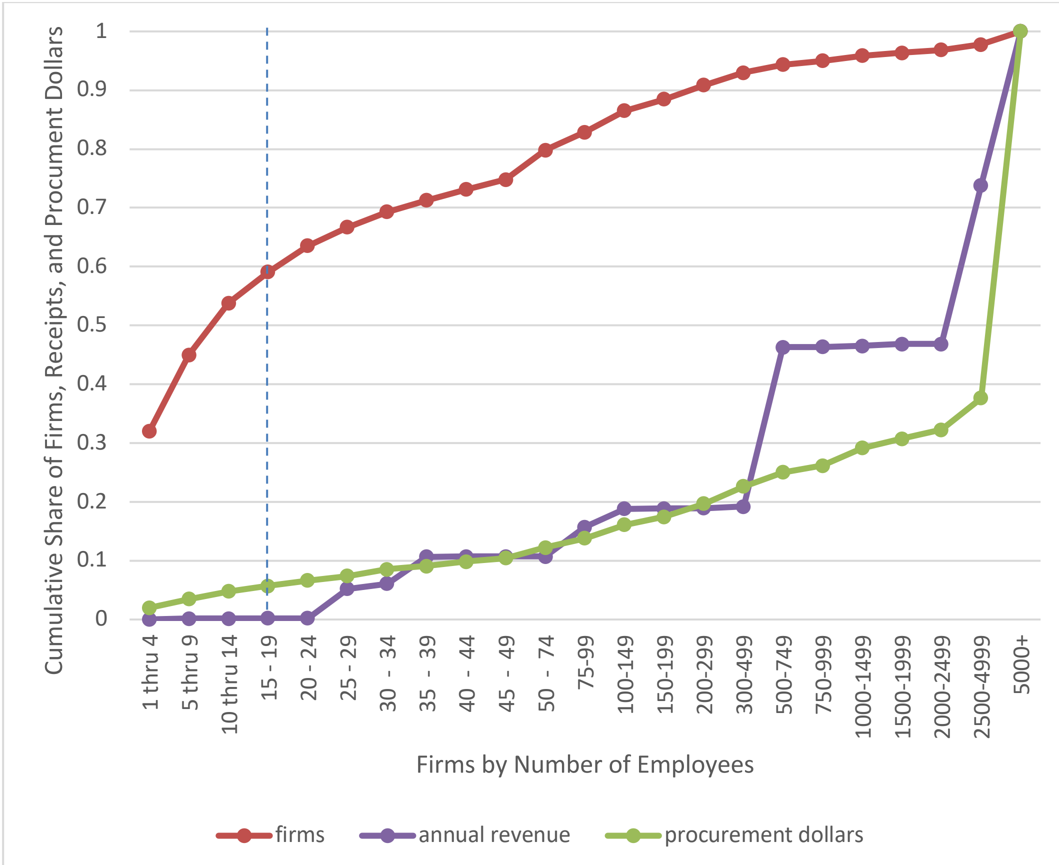


Figure 3.2. Cumulative Share of Firms, Receipts, and Procurement Dollars by Firm Size in the Procurement Economy, 2013, author's calculation.

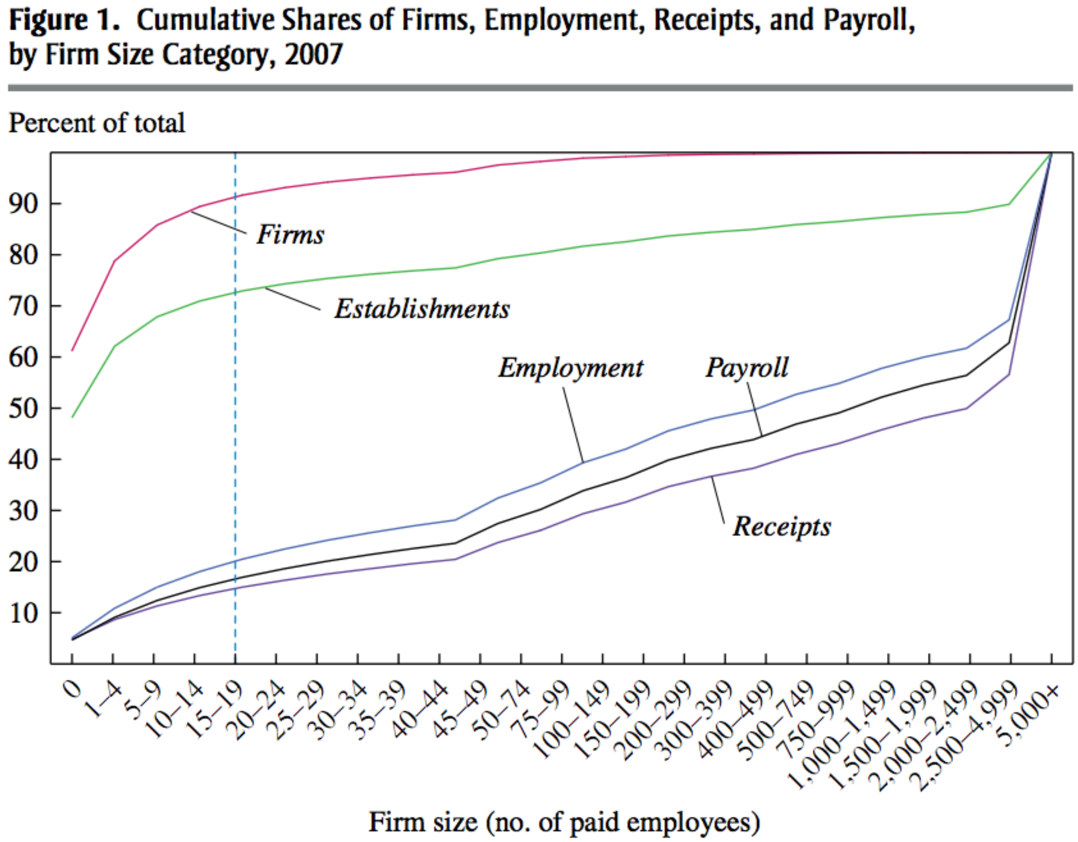


Figure 3.3. Hurst and Pugsley's (2012) Graph of Cumulative Shares of Firms, Employment, Receipts, and Payroll, by Firm Size Category, 2007.

In addition to these high-level aggregate business demographics, Hurst and Pugsley (2012) find that small firms are concentrated in a narrow range of industries and that they make up a disproportionate share of employment in those industries, implying that small firms participate in industries that have a natural small scale of production.

I follow Hurst and Pugsley's (2012) analysis by identifying small business concentration in the procurement economy at the 4-digit NAICS code industry. I find the following: First, similar to Hurst and Pugsley (2012), I confirm that small businesses in the procurement economy are concentrated in a narrow range of industries.

Second, while there is a large degree of overlap between small firm industry concentration in the U.S. economy and small firm industry concentration in the procurement economy, there are some noticeable differences. Specifically, small firms in the procurement economy participate in the manufacturing sector at much higher rates than do small firms in the broader U.S. economy. Additionally, small firm industry concentration is not correlated with small firm industry domination in the procurement economy. That is, although small firms are concentrated in a handful of industries, they are not disproportionately represented in these industries. This is a noticeable departure from Hurst and Pugsley's (2012) findings about U.S. small businesses. Finally, small procurement firms do not always specialize in one industry, suggesting that they may be specializing in something other than an industry skill. Hurst and Pugsley (2012) do not comment on industry diversification.

3.4.1 Industry Composition

I create a measure of an industry's small business concentration by quantifying the share of small businesses participating in each 4-digit NAICS code-level industry. In Fiscal Year 2013, out of the 312 possible 4-digit NAICS industry classification codes, at least one small business participated in 311 of the 312 those industry classification codes in the procurement economy.[33]

I identify the share of small firms participating in an industry, x_j, by the following:

$$x_j = \frac{s_j}{\sum_j s_j}$$

where s_j is the number of small procurement firms in industry j.

Figure 3.4 shows the cumulative share of small business participation by industry. The industry with the highest share of small businesses, x_j, is ranked 1 and the industry with the lowest share of small businesses is ranked 311.

[33] As mentioned earlier, the number of 4-digit NAICS codes varies slightly over time as the economy changes. In 2013, there were a total of 312 4-digit NAICS codes to classify all industries.

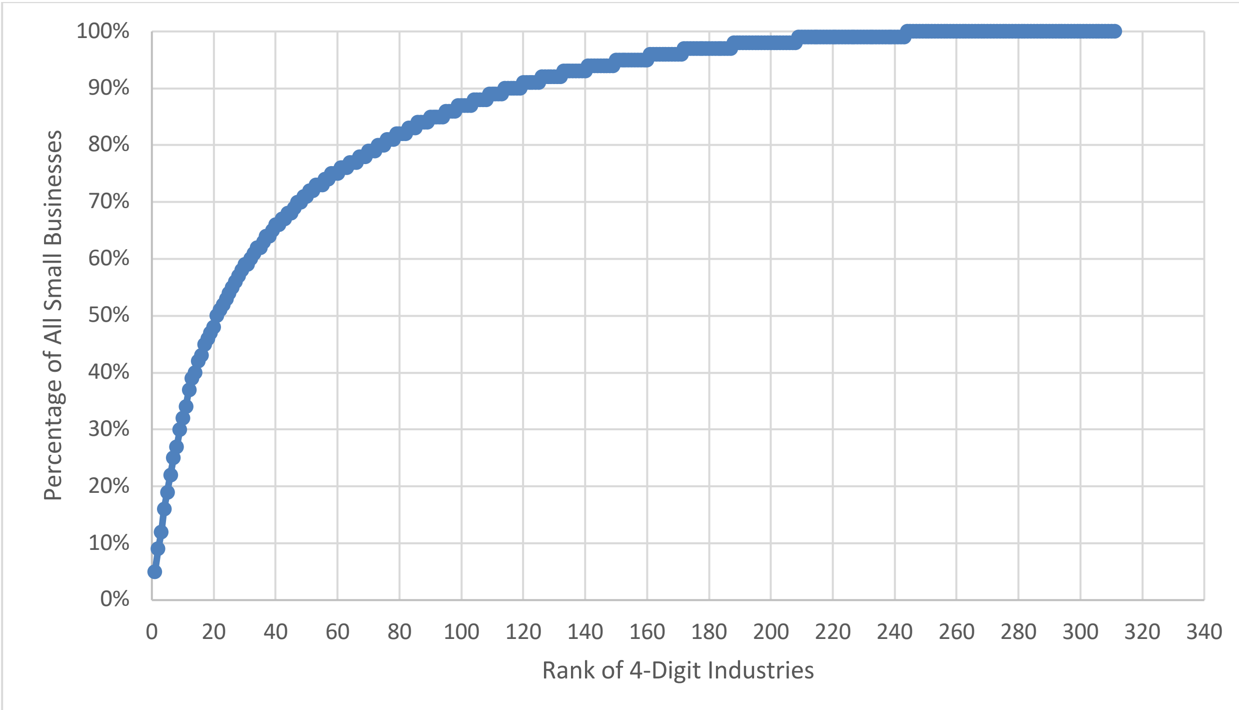
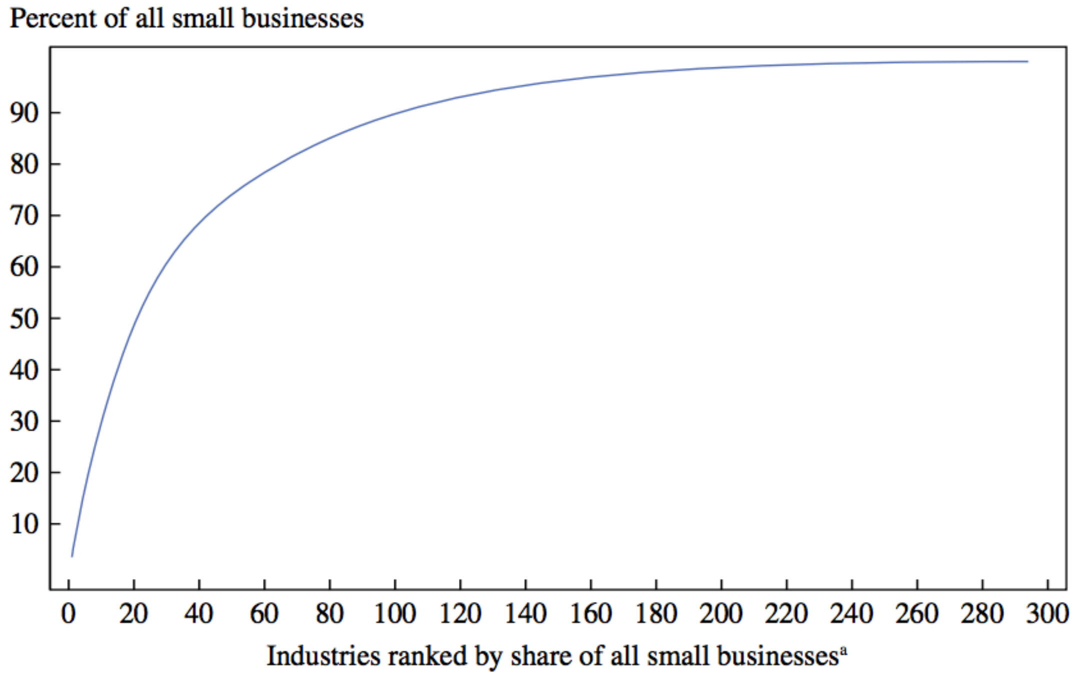


Figure 3.4. Cumulative Share of All Small Business in the Procurement Economy by Ranked 4-digit NAICS Industries, 2013, author's calculation.

Small procurement firms participate in a narrow range of industries in the procurement economy. In fact, 48% of all small procurement firms participate in just 20 industries (6% of all industries) and 65% participate in a mere 40 industries. In other words, 65% of all small procurement businesses are clustered in 13% of industries. This is almost identical to the concentration of small businesses in the regular economy. Figure 3.4 above looks almost identical to Hurst and Pugsley's (2012) cumulative industry distribution for all U.S. small businesses in 2007, reprinted below in Figure 3.5. In the regular economy, 50% of all small businesses are clustered in just 20 industries and 65% are clustered in 40 industries (Hurst & Pugsley, 2012).

Figure 2. Cumulative Share of All Small Businesses across Ranked Four-Digit Industries, 2007



Source: Authors' calculations using Statistics of U.S. Businesses data.
a. The 294 four-digit NAICS industries are ranked by their share of all businesses with fewer than 20 employees, starting with the industry with the largest share.

Figure 3.5. Hurst and Pugsley's (2012) Graph of Cumulative Share of All Small Businesses Across Ranked Four-Digit Industries, 2007.

However, if we drill in and compare the top 40 industries for small businesses in the regular and procurement economy, one similarity and two chief differences emerge between the two. Table 3.1 identifies the top 40 industries for small businesses in the procurement economy ordered by x_j. First, the procurement firm mirrors the regular economy firm in that most small procurement firms fall into the following general categories: skilled professionals (e.g., architects, consultants, computer science, physicians), skilled craftspersons (e.g., construction, electricians, industrial machinery repair), or professional service providers (e.g., clergy, real estate agents), general service providers (auto repair).

Table 3.1. Top 40 4-Digit NAICS Industries with Largest Share of Small Procurement Businesses (2013)

Rank	4-Digit NAICS Industry	Percentage of All Small Businesses	Cumulative Percent	Rank	4-Digit NAICS Industry	Percentage of All Small Businesses	Cumulative Percent
1	**Mgmt., Sci., Tech. Cons. Serv. (5416)**	5.0	5.0	21	**Legal Serv. (5411)**	1.2	49.6
2	**Arch., Eng., and Related Serv. (5413)**	3.8	8.8	22	*Other Miscellaneous Mfg (3399)*	1.1	50.7
3	**Serv. to Build. and Dwellings (5617**)	3.7	12.5	23	Vocational Rehabilitation Serv. (6243)	1.1	51.7
4	Scientific R&D Serv. (5417)	3.7	16.2	24	Software Publishers (5112)	1.0	52.8
5	**Comp. Sys. Design, Rel. Serv. (5415)**	3.1	19.3	25	Com., Ind. Mach.; Equip. Rep./Maint. (8113)	1.0	53.8
6	**Oth. Prof., Sci., and Tech. Serv. (5419)**	3.1	22.4	26	*Other Fabricated Metal Product Mfg (3329)*	1.0	54.9
7	Support Activities for Forestry (1153)	2.5	24.9	27	*Oth. Gen. Purpose Machinery Mfg (3339)*	1.0	55.8
8	**Build. Equip. Contractors (2382)**	2.4	27.3	28	**Offices of Oth. Health Pract. (6213)**	0.9	56.7
9	Business Support Serv. (5614)	2.4	29.7	29	Remed., Other Waste Mgmt Serv. (5629)	0.9	57.6
10	*Medical Eqpt and Supplies Mfg. (3391)*	2.4	32.1	30	*Aerospace Product and Parts Mfg (3364)*	0.9	58.5
11	**Nonresidential Building Con. (2362)**	2.3	34.3	31	*Communications Equipment Mfg (3342)*	0.8	59.4
12	*Nav., Electromed. Instrum. Mfg. (3345)*	2.2	36.5	32	**Lessors of Real Estate (5311)**	0.8	60.1
13	**Offices of Physicians (6211)**	2.1	38.7	33	**Religious Organizations (8131)**	0.8	60.9
14	Bus. Sch.; Comp., Mgmt. Trng. (6114)	1.6	40.3	34	**Auto Repair and Maint. (8111)**	0.8	61.7
15	Educational Support Serv. (6117)	1.4	41.7	35	**Building Finishing Contractors (2383)**	0.7	62.4
16	Prof, Com Eqpt, Suppl. Whsle. (4234)	1.4	43.2	36	RV Parks and Recreational Camps (7212)	0.7	63.0
17	Electr., Prec. Eqpt Repair, Mtce (8112)	1.4	44.5	37	Other Support Serv. (5619)	0.7	63.7
18	Investigation and Security Serv. (5616)	1.4	45.9	38	*Semicon.; Oth. Elec. Comp. Mfg (3344)*	0.6	64.3
19	**Activ. Related to Real Estate (5313)**	1.3	47.2	39	**Mach., Equip., Suppl. Whsle (4238)**	0.6	65.0
20	**Oth. Special. Trade Contractors (2389)**	1.2	48.4	40	*Oth. Elec. Eqpt and Component Mfg (3359)*	0.6	65.6

Source: USAspending.gov (2018); author's calculation. Industries that also appear on Hurst and Pugsley's (2012) top 40 industries for all small businesses in the overall U.S. economy are **bolded**. Manufacturing industries are *italicized*.

There is significant overlap between the top 40 small business industries in the U.S. economy and the top 40 small business industries in the procurement economy. As Hurst and Pugsley (2012) note, many of the top industries for all U.S. small firms in the regular economy can also be generally classified as professional services, skilled professionals, skilled craftspeople, professional services providers, and general service providers. In fact, 17 industries, or roughly 42%, appeared on both lists.

However, there are some notable differences: First, nearly a quarter of the top-ranked small business industries in the procurement economy were manufacturing related (e.g., medical equipment, electronic equipment, aerospace). *Not a single manufacturing industry broke into the top 40 for the U.S. economy.* Furthermore, Hurst and Pugsley (2012) claim the trends present in the top 40 industries persist to the top 60 industries, so it is reasonable to assume manufacturing did not break the top 60 industries for the U.S. economy. I have italicized the 9 manufacturing industries in Table 3.1 for easy identification.

Although far less important, it should also be noted that although both lists were heavily tilted toward services industries, the industries within this broad label were not always the same. For example, 12 of the top ranked small procurement business *services* industries listed were not the top service industries for all small firms in the US.

3.4.2 Robustness Checks

The results hold for alternative views of the data. For example, if I instead summarize the top 60 industries, this will represent 75% of all small procurement firms.

Additionally, 25% of these top industries remain manufacturing related. The most notable difference when analyzing the top 60 industries instead of the top 40 is that there are a disproportionate number of construction-related industries in the 40-60 rank range, so if anything, the results skew away from the more purely professional services and toward construction and manufacturing. Similar to Hurst and Pugsley (2012), I excluded nonemployee firms from analysis, although results were similar.

The procurement economy is not a perfect microcosm of the regular economy, so it is also possible that these results are skewed in that more small firms are concentrated in an industry simply because the industry itself is dominant. That is, perhaps the industries with the highest share of small business are simply the largest industries in the procurement economy, so it is natural that the industries that have the highest concentration of all businesses would also have the highest concentration of small businesses.

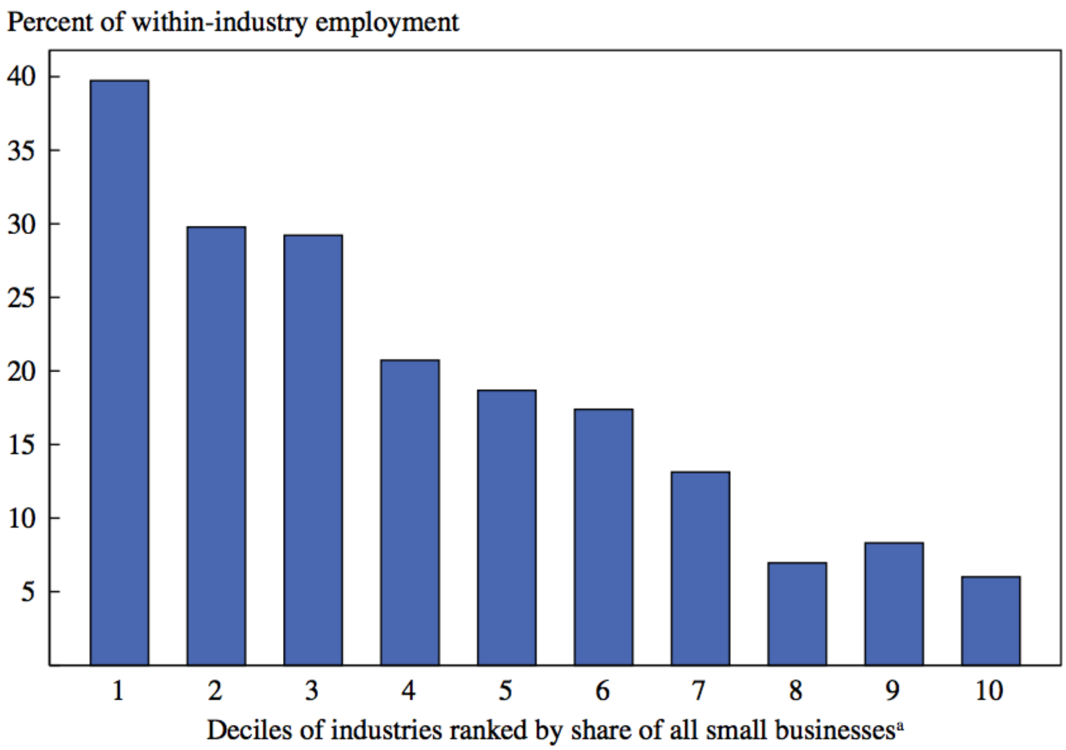
Similar to Hurst and Pugsley (2012), I perform the following robustness check by creating a weighted measure of rank, $\widetilde{x_j}$, by controlling for the relative proportion of businesses in the industry:

$$\widetilde{x_j} = a_0 + a_1\left(\frac{n_j}{\sum_j n_j}\right) + \varepsilon_j$$

where n_j is the number of businesses in j industry regardless of size. Table C2 in Appendix C provides the top 40 adjusted-rank industries for small businesses in the procurement economy. The results look very similar: 35 of the top 40 ranked industries remain in the top 40 and the types of industries are very similar. Nearly a quarter of the top 40 industries for small businesses in the procurement economy are in the manufacturing sector and the remaining industries fall into the broadly defined service industry categories listed above.

Although small procurement firms are concentrated in a handful of industries much like the U.S. economy, they are not performing the majority of work in these narrow industries. Hurst and Pugsley (2012) find a statistically significant relationship between those industries with the highest concentration of small firms with those industries with the highest share of employment in small firms. In other words, in the overall U.S. economy, small firms tend to concentrate in industries where the natural scale is the small firm. Hurst and Pugsley's graph is reprinted for comparison in Figure 3.6.

Figure 3. Small Business Share of Within-Industry Employment, by Decile of Ranked Four-Digit Industries



Source: Authors' calculations using Statistics of U.S. Businesses data.

a. The 294 four-digit NAICS industries in figure 2 are grouped into deciles. Reported percentages are simple averages for the industries in the indicated decile.

Figure 3.6. Hurst and Pugsley's (2012) Graph of Small Business Share of Within-Industry Employment, by Decile or Ranked Four-Digit Industries.

Unfortunately, I could not produce a similar comparison because I do not have data on the number of employees who participate in the procurement economy for large firms. However, instead of using employment, I use procurement contract dollars as the metric of comparison. That is, I analyzed the relationship between the industries with the highest concentration of small firms and the industries' share of small firm contract dollars. To measure this, I create y_j, industry j's share of contract dollars performed by small businesses.

$$y_j = \frac{r_j^s}{r_j^n}$$

where r_j^s is the procurement revenue earned by small businesses in industry j and r_j^n is the total procurement revenue in industry j.

I then group and average these shares by industry deciles as shown in Figure 3.7. Each decile contains one-tenth of all industries, or 31 of 311 industries in the set. Each bar in Figure 3.7 represents 31 industries, or a tenth of all 4-digit NAICS industries. The first decile bar, labeled 1 in Figure 37, contains the average share of revenue of 31 industries with the *highest share of small business participating in the procurement economy*. The y-axis identifies the industry share, by contract dollars, that these small firms have averaged across the NAICS codes in that decile.

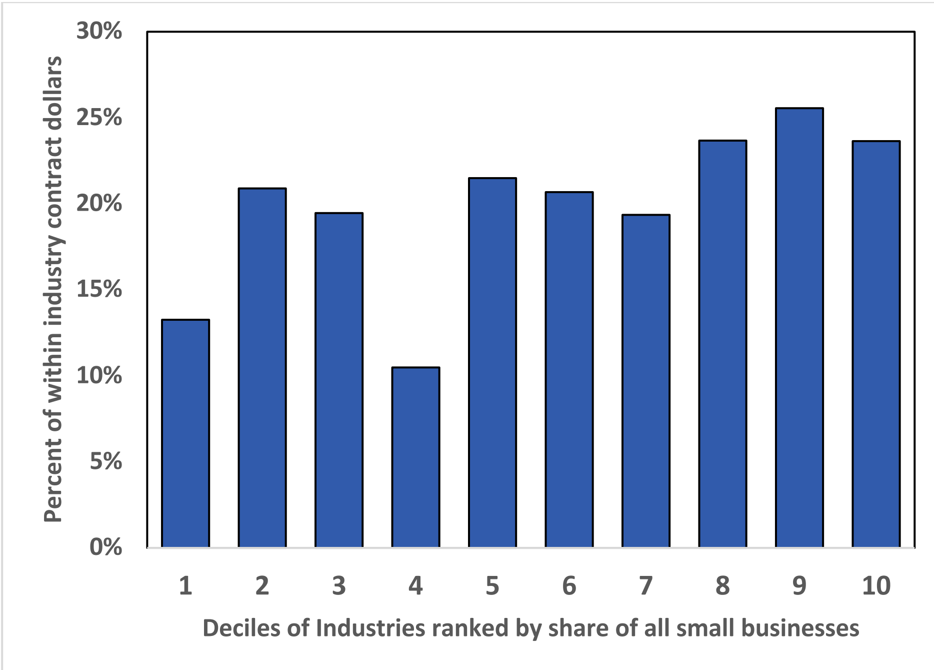


Figure 3.7. Small Business Share of Within-Industry Contract Dollars, by Decile of Ranked Four-Digit Industries. The first decile has the most number of participating small businesses. Author's calculation using USAspending.gov (2013) data.

In the procurement economy, there does not appear to be any relationship, positive or negative, between industries whose concentration of small firms is greatest and the industries' share of contract dollars awarded to small firms. The *p*-value for the difference between the average of decile 1 and 2 is .10, but even if it were statistically significant, the predicted difference is in the wrong direction! Here, decile 2 is earning a higher share of contract dollars than decile 1. There is a statistically significant difference between decile 3 and decile 4, but this seems less noteworthy given that the later deciles, deciles 5-10, have an even larger concentration of small firms than do deciles 1-3. Perhaps the most striking finding from Figure 3.7 is that there does not seem to be any natural concentration of small businesses in the procurement economy as measured by contract dollars, despite an overwhelmingly large concentration of small firms in the first two deciles.

In fact, if we just compare the first 5 deciles, deciles 1-5, to the last 5 deciles, deciles 6-10, there is a statistically significant difference between the two means. The last 5 deciles, those deciles with the lowest concentration of small firms, have a statistically significant higher share of small business dollars (23%) than the first 5 deciles (17%).[34] A somewhat comparable graph for the overall economy profiled above had a strong negative relationship; but here, if anything, the relationship is positive. If I instead average all small business dollars by decile rather than take the average of the averages of each industry within the decile, I get similar results. See Appendix for figure.

In the overall economy, small firms are concentrated in a narrow range of industries and these industries are also overrepresented by small firm employment, suggesting that these industries are amenable to small-scale operations. In the procurement economy, small firms are concentrated in a narrow range of industries, but these industries do not seem to be overrepresented by small firms. Instead, the industries in which small procurement firms cluster appear to be industries with a large share of overall contract dollars. That is, in the procurement economy, small firms are concentrated in industries with a large share of total contract dollars rather than

---

[34] T-stat is -2.15.

concentrated in industries with a large share of small firm contract dollars (or employment).

3.5 Small Business Success in the Procurement Economy

What is correlated with firm success in the procurement economy? I test whether a small subset of young firms in high-growth-related industries are more likely to grow in the procurement economy than mature, small firms. Finally, I develop a proxy for procurement specialization and test whether this predicts success.

I isolate U.S.-based small firms which registered to do business with the federal government in fiscal year 2013[35] and won at least one contract in fiscal year 2013.[36] I track which firm and contract characteristics are correlated with firm success in the procurement economy, with firm success imperfectly measured by federal contract dollars obligated from fiscal years 2013-2017. I use contract dollars per employee as an alternate metric of success. For simplicity, I remove any firm that registered in fiscal year 2013 and was bought during this time period. I also limit the analysis to firms with greater than 0 employees and fewer than 11 employees in FY2013. I verified employment data with another database, Mergent Intellect. All firms in my dataset also

[35] I only included firms which registered after the System for Award Management (SAM) migrated platforms. I randomly sampled firms in my dataset to confirm that this was accurate. No firms in my sample were registered before 2013.
[36] These firms were awarded at least one contract action with positive dollar amounts, so this is already selecting a more successful class of firms than the full universe of those which registered to do business with the federal government but never won a contract.

qualify as small business for at least one NAICS code. There are 3,088 firms in my sample.

First, I control for firm age. This is distinct from contractor registration date. Table 3.2 shows the distribution of firms in my sample, the majority of which are new firms.

Table 3.2. Firm Age of Small Firms Registering in the System for Award Management (SAM) in FY2013

Firm Age	Frequency	Percentage
0-5	2,036	65.93
6-10 years	378	12.24
11-15 years	162	5.25
16-20	117	3.79
21+	395	12.79
Total	3,088	100

For contract-specific variables, I test whether it matters what the firm provides and to whom. Specifically, I test whether industry specialization affects small firm success as defined by federal contract dollars obligated. To do that, I include the following industry-related metrics: Is participation in a particular 2-digit NAICS code industry—a top 40 small business industry as outlined in Sections 3.1-3.4, or the high growth industries similar to what Haltiwanger et al. (2010) find in the regular economy—correlated with firm success? The list of high-growth industries is in Appendix C, Table C3.

Finally, I extend the test of firm specialization to test whether firm diversification, as measured by 4-digit NAICS codes, is correlated with firm success, but I also explore what firm diversification means. Table 3.3 lists the number of firms by how many distinct NAICS codes in which a firm was awarded a contract.

Table 3.3. Number of Firms by Number of Different NAICS Codes in Which Firms Are Contracting

Number of Different 4-digit NAICS Codes of Contracts Awarded	Frequency	Percentage
1	2,377	76.98
2	462	14.96
3	119	3.85
4 or more	130	4.18

Note. *N* = 3,088.

If winning contracts in multiple NAICS codes is simply the result of coding error, first we should expect to see the significance of the number of contracts increase if we remove NAICS codes from the regression. That is, the more contracts a firm wins, the more likely it is to win in multiple NAICS codes, and therefore removing NAICS codes from the regression should strengthen the significance of number of contracts won. I test for this and include results in Appendix C, Table C4.

Second, if it is simply coding error, firms should **not** be *self-reporting* that they have expertise in multiple 4-digit NAICS code. I test this assumption using data from the System for Award Management (SAM), the federal government's repository of all businesses that wish to contract with the federal government. When registering, firms are required to identify their primary 6-digit NAICS code, but they can specify as many NAICS codes as they like. Because identifying this data is onerous, I sampled self-reported responses in SAM regarding NAICS code specialization of the 100 worst and 100 best performing firms in my dataset at the time of registration. The average number of NAICS codes that the worst performing small businesses reported was 1.21. The best 100 firms identified on average 3.86 4-digit NAICS codes.[37]

Self-reporting specialization in multiple 4-digit NAICS codes means that either firms are actually proficient in all of these NAICS codes or they understand that diversification is necessary in the procurement economy for authentic or inauthentic reasons. If small firm success in federal procurement requires *either* authentic diversification in multiple industries (as imperfectly measured by NAICS codes) or inauthentic diversification of knowing to search for and bid on misclassified contracts, then self-reported proficiency in multiple NAICS codes can be construed as a proxy for understanding how the procurement world works. That is, perhaps small firm success is correlated more with procurement economy knowledge rather than specialization in a particular industry. In other words, it is more important to know how to write a successful contract proposal than to have deep expertise in a specific industry.

[37] To be clear, they reported far more 6-digit NAICS codes, but I grouped them at the 4-digit level.

Therefore, doing business in multiple NAICS codes is not necessarily reflective of specialization in those NAICS codes per se but could actually be reflective of specific federal procurement knowledge or specialization in procurement rules and regulation. Diversification may be extremely successful in the federal procurement market. However, for that to be true, it would mean that firms with fewer than 11 employees are managing to diversify in multiple 4-digit NAICS codes and traverse the federal procurement rules and regulations. This seems implausible. It seems even more implausible to assume that these small firms are also simultaneously participating in the regular economy.

Mergent Online (MergentOnline.com) provides suggestive evidence that self-reporting specialization in a number of NAICS codes in SAM is signaling understanding of how the procurement economy works rather than authentic diversification. Mergent Online, a database that has historical data on all US firms, also reports a firm's primary and secondary 6-digit NAICS code. Not a single firm in my dataset reported more than one 6-digit NAICS code in Mergent Online, which identifies primary and secondary 6-digit NAICS codes based on company business description in business filings.[38] This means that an ostensibly objective outsider rating a firm's capabilities by its business filing fails to identify more than one line of business for any small firm in my dataset. Yet, there is a large discrepancy between the most and least successful firms' *self-*

[38] Personal communication with an employee at Mergent Online.

reported lines of business in a government database. And the differences between the Mergent Online reporting and SAM reporting would be more exaggerated if I were reporting SAM NAICS codes at the 6-digit level rather than the 4-digit level as Mergent Online does.

The most plausible explanation is that the number of NAICS codes registered by a firm is actually evidence of its understanding the procurement economy and its regulations rather than specializing in a particular industry. Furthermore, understanding procurement regulations and processes represents a substantial fixed cost for a firm, small or large. If the firm owner already has this knowledge because of a former government role, it represents an enormous advantage over those who need to invest in acquiring this expertise.

Finally, I control for additional procurement-specific characteristics: I include a dummy variable for whether a firm wins contracts from the Department of Defense (DOD), given that DOD has the largest contract budget. Roughly 29% of firms (885) in my sample won at least one contract with DOD. I then also test whether it matters if the firm is able to win contracts from different agencies or different subagencies. Almost 10% of firms in my sample won contracts at two or more agencies. Finally, I also test whether the number of contracts won affects firm success.

I also control for firms with special bid preference status within the procurement economy. Firms can certify that they have one or more of the following owner characteristics: minority owned[39], woman owned, Native American/Alaskan native owned, etc. Contracts may be set-aside (restricted to) for one or more of these classifications. The goal of the federal government is to set aside 23% of all contract dollars for one or more of these classifications[40]. In my analysis, I call this variable special status and treat it as binary: Firms with at least one special firm designation are classified as 1, or classified as 0 if they do not have at least one designation; 47% of firms in my sample had at least one special firm designation.

3.5.1 Specification

I perform the following linear regression:

$$\text{Log}[Total\ contract\ dollars] = \alpha + b_1 no.of\ fy17\ employees + b_2 bidpreference + b_3\ no.of\ NAICS\ codes + b_4 agencies + b_5 subagencies + b_6 no.of\ contracts + b_7 DOD + b_8 industry + b_9 state + b_{10} firm\ age + b_{11} top40sb\ industry + \varepsilon$$

I use both FY13 employee count and FY17 employee count and neither changed results.

---

[39] 8(a) firms and small disadvantaged firms are likely to be minority-owned, but the test is "socially & economically disadvantaged" – where the hurdle for proving social disadvantage is met if the firm is owned & controlled by one or more: African Americans, Hispanic Americans, Native Americans, Subcontinent Asian-Americans, or Asian-Pacific Americans.

[40] https://www.sba.gov/sites/default/files/resources_articles/FY19_Small_Business_Goaling_Guidelines_Draft_2018-08_Final.pdf

3.5.2 Results

Results are reported in Table 3.4. Standard errors are in parenthesis. The reference firm age category is 0-5 years or older. The full regression with all other variables is included in Appendix C, Table C5.

Table 3.4. Linear regression; response Variable is Log[*Total Contract Dollars*]

	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
No. of NAICS	0.302***	0.229***	0.214***	0.214***	0.206***	0.206***
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
No. of agencies		0.483***	0.443***	0.439***	0.439***	0.436***
		(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
DOD contract			0.474***	0.487***	0.487***	0.486***
			(0.07)	(0.08)	(0.07)	(0.07)
High-growth industry				0.119	0.128*	0.126*
				(0.06)	(0.06)	(0.06)
6-10 years					-0.624***	-0.630***
					(0.11)	(0.11)
11-15 years					-0.707***	-0.713***
					(0.12)	(0.12)
16-20 years					-0.623***	-0.635***
					(0.08)	(0.08)
21+ years					-0.402***	-0.412***
					(0.09)	(0.09)
fy17emp						0.005
						(0.01)
_cons	9.633***	9.189***	9.122***	9.088***	8.78***	8.863***
	(0.07)	(0.10)	(0.09)	(0.09)	(0.12)	(0.12)
N	3088.000	3088.000	3088.000	3088.000	3088.000	3088.000
vce	robust	robust	robust	robust	robust	robust

Note. Standard errors are in parenthesis. The reference firm age category is 0-5 years or older.

Being a young firm, from 0-5 years old, is strongly correlated with firm success: Relative to mature firms, being young predicts a 40% increase in contract dollars obligated over the fiscal years 2013-2017.

Although not a perfect apples-to-apples comparison, this correlation between young firm age and success appears to contradict lifecycle dynamics in the regular economy. Figure 3.8 plots average log contract dollars by firm age. Note that this does not include dollars (revenue) earned outside of the procurement economy.

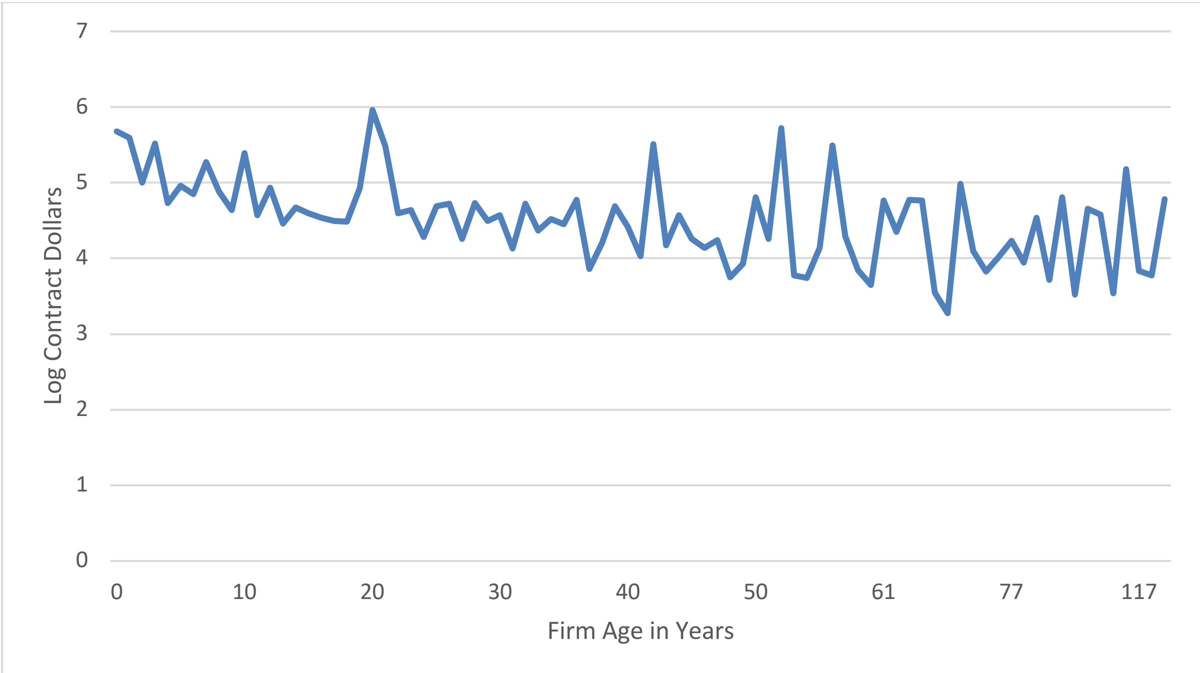


Figure 3.8. Sample's FY13-17 Average Log(Contract Dollars) by Firm Age in Procurement Economy.

Figure 3.9 is a copy of Kueng et al.'s (2014) graph plotting log revenue by firm age for a random stratified sample of 6,500 Canadian firms tracked from 1996-2006.

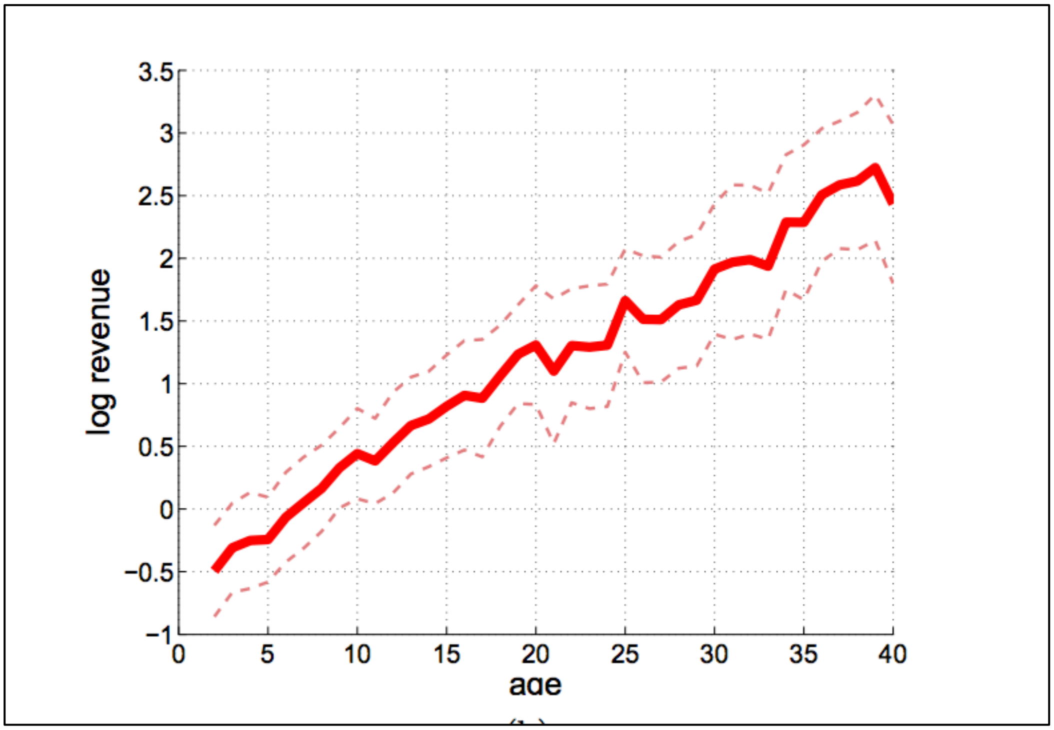


Figure 3.9. Kueng et al.'s (2014) Graph Plotting Log Revenue by Firm Age for a Random Stratified Sample of 6,500 Canadian Firms Tracked From 1996-2006.

Second, in contrast to the regular economy, specializing in high-growth industries does not seem to affect firm success in the procurement economy, as measured by contract dollars. Third, although not featured in the above regression because of space constraints, I controlled for 2-digit NAICS codes and this too did not affect results. Furthermore, there appears to be no statistical significance if one's primary 4-digit NAICS code is one of the top 40 NAICS codes for small businesses. I tested this with the top 40 industries in the overall economy and in the procurement economy. Although the regression differs from Hurst and Pugsley's (2012) and Haltiwanger et al.'s (2016) models for young firm success, they too found that being in one of the top 40 small business-intensive industries did not lead to firm success, as measured by an increase in employment.

However, firm diversification, as measured by the number of distinct NAICS codes in which a small firm wins contracts, is strongly correlated with firm success. Specifically, for every additional contract with a distinct 4-digit NAICS code, federal dollars obligated increases by 48%. Diversification seems to be beneficial to small firms in the procurement sector. Although not featured in the above regression because of space constraints, this correlation between number of NAICS codes and contract dollars holds, even when controlling for the number of contracts won.[41]

Not surprisingly, having at least one contract with DOD is by far the largest predictor of firm success: DOD awards roughly 60%[42] of all contract dollars in any fiscal year. Winning at least one contract with the Department of Defense is strongly positively correlated with contract dollars won during fiscal years 2013-2017. In fact, one win increases contract dollars obligated by almost 80%. Diversifying with other agencies also indicates success, although diversification within an agency at the subagency level is not statistically significant. Controlling for contracts awarded, for every additional agency with which the firm does business, there was a correlated increase in awarded dollars by 50%. Thus, while it is more important to be able to win a

---

[41] While the number of contracts is positively correlated with contract dollars earned, it is fairly small. Every additional contract won increases contract dollars obligated by .001 percent. See Appendix C.
[42] For example, in FY18, DOD contract spending accounted for 65% of all contract spending (FPDS-NG, n.d.); author's calculation.

contract with DOD, the most successful small firms can work with DOD and civilian clients.

3.5.3 Robustness Checks

Although I controlled for employment size in the above regression, I ensured that the importance of doing business in multiple NAICS codes held even after controlling for dollars per worker. To do this, I ran the same regression as above but changed the response variable to the log of contract dollars per employee. I tested this relationship with FY13 and FY17 employment data and the results were similar to each other and to the original regression in Table 3.4 above. Table 3.5 shows the regression using FY17 employment data.

Table 3.5. Regression Where Response Variable is Log[*Total Contract Dollars/FY17 Employee*]

	(1) b/se	(2) b/se	(3) b/se	(4) b/se	(5) b/se	(6) b/se
No. of NAICS	0.273***	0.223***	0.213***	0.214***	0.195***	0.195***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
No. of agencies		0.335***	0.308***	0.306***	0.301***	0.298***
		(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
DOD contract			0.326***	0.322***	0.314***	0.326***
			(0.08)	(0.08)	(0.08)	(0.08)
Special status				-0.088	-0.009	-0.009
				(0.07)	(0.06)	(0.06)
6-10 years					-1.118***	-1.114***
					(0.14)	(0.14)
11-15 years					-1.265***	-1.281***
					(0.14)	(0.14)
16-20 years					1.469***	-1.466***
					(0.09)	(0.09)
21+ years					-0.860***	-0.865***
					(0.10)	(0.10)
High-growth industry						0.112
						(0.07)
_cons	9.026***	8.714***	8.666***	8.629***	9.127***	9.095***
	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
N	2791.000	2791.000	2791.000	2791.000	2791.000	2791.000
vce	robust	robust	robust	robust	robust	robust

Note. *N* is slightly smaller in this regression (2,791) than in the original analysis (3,088) because in FY17, some firms had zero employees and were excluded from the regression because the response variable was undefined.

The only difference is that with this new response variable, being in the top 40 industries in the regular economy or the procurement economy is correlated with firm success. Although being in the top 40 does not affect the strength or size of the other variables, it too is a statistically significant variable.

Finally, for simplicity, I include a simple scatter plot between log contract dollars per FY17 employee and the number of distinct NAICS codes in which the firm won contracts (Figure 3.10). There is a clear positive relationship between the two variables.

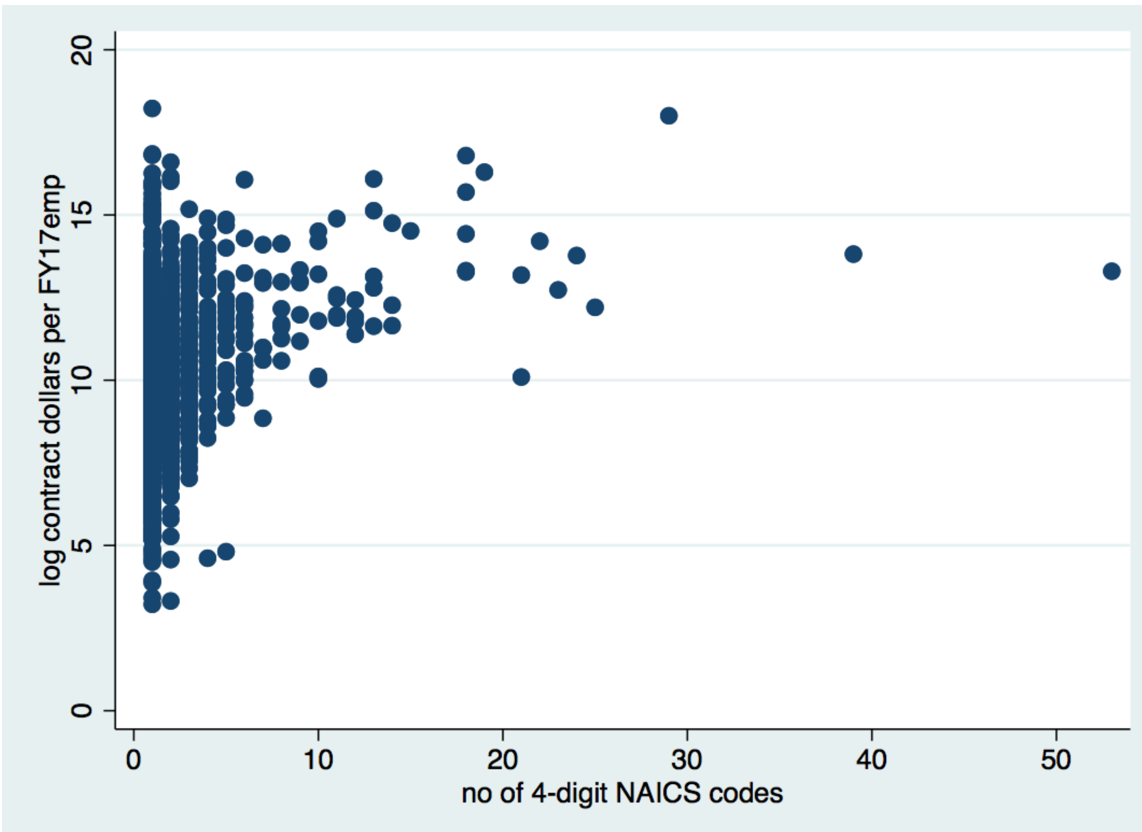


Figure 3.10. Simple Scatterplot of Log Contract Dollars per FY17 Employees vs Number of NAICS Codes.

3.6 Closing

Although more research is needed, these findings suggest that small procurement firms may have different growth patterns than small and new businesses in the larger U.S. economy. First, while they are clustered in a narrow subset of industries much like all small firms, small procurement firms participate in different industries. Second, these industries are strongly correlated with the composition of the procurement economy: 65% of small procurement firms are clustered in the 40 industries that make up roughly 65% of the procurement economy. This too is a departure from small firm concentration in the regular economy where small firms are concentrated in industries that are also overrepresented by small firm employment. Third, small businesses in the procurement economy are overrepresented in manufacturing industries relative to the regular economy.

Fourth, and most importantly, these findings suggest a small procurement firm's success may be the result of a factor other than simply excelling in a particular sector or industry. Given that the most successful firms in the procurement economy seem to be performing work in several different industries—a fact that seems less common in the overall economy—their success seems to be related to something other than excellence in a particular specialty.

Moreover, these findings have several important implications for future research and policy. First, the research presented here suggests that small and new businesses participating in the procurement economy have different growth patterns than other U.S. firms and should be studied separately from other business demographics. *Inc. Magazine's* identification of its fastest growing firms actually segments them into government procurement as its own category[43]; researchers may want to differentiate similarly.

Additionally, this research suggests that mastery of procurement regulations and processes[44] may be a main predictor of firm success in the procurement economy. Can registration of a number of NAICS codes predict federal procurement success for small firms? Or are there other proxies for insider federal procurement knowledge? Can we test whether this predicts firm success in the procurement market? For example, are there responses in SAM registration, such as number of NAICS codes selected, that clearly demonstrate familiarity with the federal procurement market, and can we test to see if these predict firm success in the procurement market?

Finally, are these findings regarding small firm dynamics in the federal procurement economy generalizable to state procurement and local procurement ecosystems? That is, is procurement specialization more important than industry expertise, and should small firms doing business at the state and local levels be studied separately from other small firms? And does this also extend to large firms in the procurement economy? At present, we lack reliable data to explore this issue. However,

---

[43] See https://www.inc.com/inc5000/2019/top-private-companies-2019-inc5000.html

[44] Procurement regulation is the FAR and its supplements, agency manuals, etc. For example, a firm doing business with DHS must navigate the FAR, HSAR (DHS FAR Supplement) and HSAM (Homeland Security Acquisition Manual).

given that federal, state, and local procurement is roughly $1.5 trillion of the U.S. economy, this issue warrants further study.

APPENDIX A. BETA REGRESSION ROBUSTNESS CHECK

Table A1. Beta Regression Robustness Test

% approved	(i)	(ii)
Male	0.12***	0.12***
Republican	-0.09*	-0.19***
Independent	-0.35***	-0.23***
Male # Republican	-0.14**	
Male # Independent	0.16**	
age	0.01***	0.01***
cases	0.00***	0.00***
N	2517	2517
r2_a		

legend: * p<0.05; ** p<0.01; *** p<0.001

APPENDIX B. VARIED DEFINITION OF RUSHED JUDGES

Table B1. Varying the Definition of Rushed From an Increase of 30% to an Increase of 25% or 35%

	(1)	(2)
rush25percent	0.0594	
	(1.86)	
rush35percent		0.0899*
		(2.37)
gender	0.101***	0.102***
	(4.84)	(4.89)
Republican	-0.167***	-0.167***
	(-7.01)	(-7.00)
Independent	-0.204***	-0.204***
	(-8.11)	(-8.12)
age	0.0113***	0.0112***
	(10.18)	(10.13)
Cases	0.000449***	0.000459***
	(3.49)	(3.53)
N	2377	2377

t statistics in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

APPENDIX C. SUPPLEMENTS TO CHAPTER 3

Method for Unit of Analysis of Firm and Establishment Using DUNS Number and ParentDUNS Number as a Proxy

USAspending.gov tracks data at the DUNS and ParentDUNS level. Firms can have several DUNS numbers. DUNS numbers must be obtained for each location of a business, so the DUNS number appears to be most similar to an establishment although not identical (D-U-N-S number, n.d.).[45] ParentDUNS is defined as the DUNS number's immediate headquarters or parent company (HQ/Parent D-U-N-S number, n.d.), which appears to be closer to the definition of a firm although not identical. When possible, I used ParentDUNS instead of DUNS. To identify small firms with fewer than 20 employees for my analysis of industrial composition in section 3.4 of Chapter 3, I removed any ParentDUNS number, and by extension any DUNS number, that listed a number of employees count larger than 19 at any point during the year. This is by no means a perfect method. One serious drawback of this method is that I count a large firm as a small firm if all of the locations (or establishments) of the large firm or its establishments have fewer than 20 employees.

---

[45] An outdated document outlining what DUNS is was formerly available at https://procurement.inl.gov/Small-Business-Program/Shared%20Documents/DUNS%20FAQ.pdf.

I applied more robustness checks when analyzing drivers of small firm success. There were roughly 3,850 firms originally identified as earning a non-zero amount of FY13 contract dollars and having 1-10 employees. I cross-checked this data with another database, Mergent Online, and eliminated any firms from my sample that had fewer than 1 employee or greater than 10 employees in all of their establishments combined. I eliminated roughly 700 firms from my sample, although the majority of firms I eliminated either listed 0 employees or roughly 11-20 employees, rather than having a very large number of employees. This suggests that some of the discrepancies were merely a product of when the results were reported in the year rather than underlying misclassification. Furthermore, I ran the regression including all firms from the original sample and got virtually the same results as the winnowed-down sample.

Definition of Small Business in the Federal Procurement Economy

Although the federal government has its own definition of small business for the purposes of procurement, I will use Hurst and Pugsley's (2012) definition of a *small business* as fewer than 20 employees, unless otherwise noted, so that results of small business industrial composition in the procurement economy are more comparable to Hurst and Pugsley's results. My analysis excludes nonemployer firms but results are similar when they are included. Note that this classification differs from what the federal government defines as small business.

In order to ensure that the government awards 23% of federal procurement dollars to small business, it must first be able to identify a small business. The U.S. Small Business Administration (SBA),[46] as manager of the program, creates the small business definition. Section 3 of the original Small Business Act of 1953 states that "a small business concern shall be deemed to be one which is independently owned and operated and which is not dominant in its field of operation".[47] The SBA has translated this qualitative definition into industry-specific thresholds, usually based on a firm's annual revenue or employee count; these thresholds have evolved over time to adapt to changes in an industry's market structure or inflation.[48] These definitions are extremely complex, usually much larger than other developed countries' and the American public's small business definitions, and are subject to interest-group tinkering.

The SBA identifies small-business size thresholds by North American Industrial Classification System (NAICS) code. The NAICS, replacing an earlier Standard Industrial Classification (SIC) system in 1997, is a classification system describing a firm's industry. There are over 300 different 6-digit NAICS codes, each representing a specific subsector of economic activity and rolled into one of 25 broad 2-digit NAICS sectors, such as

---

[46] The U.S. SBA, with varying degrees of direction from Congress. The Small Business Act 3(a)(1) caps many agricultural-industry size thresholds at $750,000 in annual revenue.

[47] 15USC Subsection 632(A)(1), https://www.law.cornell.edu/uscode/text/15/632

[48] The U.S. SBA (2014) provides a list of the most current small business thresholds by 6-digit NAICS code.

Manufacturing or Information. The SBA treats each of these 6-digit NAICS codes as its own industry, potentially warranting a distinct size business threshold.

Today, a set of 26 employee- and annual-revenue-based thresholds mapped at the 6-digit NAICS code level create the jagged dividing line between small and large business.[49] Table C1 provides a summary of the highest and lowest thresholds by employee size and annual revenue and includes the number of industries defined by these thresholds. In general, the SBA assigns employee-based size restrictions for manufacturing industries and revenue-related size restrictions for services industries.

Table C1. U.S. Small Business Administration Threshold Ranges for Business Size

Type of Threshold	Small Business Threshold Ranges	
	Low*	**High**
Annual revenue	$5.5M (4 industries)	$38.5M (59 industries)
Number of Employees	50 (1 industry)	1,500 (14 industries)

* The SBA does not have the authority to set agriculture-related NAICS codes thresholds. These are determined by statute and currently capped at $750K.

[49] For more information on the SBA's methodology, see Cornell Law School's Legal Information Institute (n.d.).

Table C2. Adjusted Rank of Percentage of Small Procurement Businesses in Top 4-Digit NAICS Industries

Rank	4-Digit NAICS Industry	Percentage of All Small Businesses	Cumulative Percentage	Rank	4-Digit NAICS Industry	Percentage of All Small Businesses	Cumulative Percentage
1	**Mgmt, Scientific, and Tech Cons Serv. (5416)**	4.4	4.4	21	*Other General Purpose Machinery Mfg (3339)*	1.2	47.6
2	**Architectural, Engineer, and Related Serv. (5413)**	4.2	8.7	22	*Aerospace Product and Parts Mfg (3364)*	1.1	48.7
3	**Comp. Systems Design and Related Serv. (5415)**	3.7	12.4	23	*Other Misc Manufacturing (3399)*	1.1	49.8
4	Scientific R&D Serv. (5417)	3.4	15.8	24	Comm., Ind. Mach., Equip. Repair, Maint.(8113)	1.1	50.9
5	**Serv. to Buildings and Dwellings (5617)**	3.3	19.1	25	Educational Support Serv. (6117)	1.1	52
6	**Nonresidential Building Construction (2362)**	2.7	21.8	26	**Other Specialty Trade Contractors (2389)**	1	53
7	**Oth. Prof, Scientific, Tech. Serv. (5419)**	2.6	24.4	27	Nursing Care Facilities (6231)	1	54
8	**Building Equipment Contractors (2382**)	2.5	26.9	28	Remediation, Other Waste Mgmt Serv. (5629)	1	54.9
9	*Nav., Electromed., Control Instrum. Mfg (3345)*	2.4	29.4	29	**Legal Serv. (5411)**	1	55.9
10	*Medical Equipment and Supplies Mfg (3391)*	2.3	31.7	30	*Comm Equip. Manufacturing (3342)*	0.9	56.8
11	Support Activities for Forestry (1153)	1.8	33.5	31	**Activities Related to Real Estate (5313)**	0.9	57.7
12	Business Support Serv. (5614)	1.7	35.2	32	*Semiconduct.., Oth. Elec. Component Mfg. (3344)*	0.8	58.5
13	**Offices of Physicians (6211)**	1.7	36.8	33	Facilities Support Serv. (5612)	0.7	59.3
14	Traveler Accommodation (7211)	1.6	38.5	34	**Lessors of Real Estate (5311)**	0.7	60
15	Investigation and Security Serv. (5616)	1.4	39.9	35	Machine Equip. and Suppl. Merch. Whsle. (4238)	0.7	60.7
16	Prof., Comm. Equip Suppl. Merch. Whsle. (4234)	1.4	41.3	36	**Auto. Repair and Maint. (8111)**	0.7	61.4
17	Elec./ Prec. Equip. Repair and Maint. (8112)	1.4	42.7	37	Vocational Rehabilitation Serv. (6243)	0.7	62.1
18	Bus. Schools, Comp., Mgmt Training (6114)	1.3	43.9	38	Other Heavy and Civil Engineering Con (2379)	0.7	62.7
19	Software Publishers (5112)	1.2	45.2	39	*Oth. Elec. Equip. and Component Mfg (3359)*	0.7	63.4
20	*Oth. Fabricated Metal Prod. Mfg. (3329)*	1.2	46.4	40	Highway, Street, and Bridge Con. (2373)	0.6	64.1

Note. Does not include zero-employee firms. USAspending.gov (2013); author's calculation. Industries that also appear on Hurst and Pugsley's (2012) top 40 industries for all small businesses in the overall U.S. economy are bolded. Manufacturing industries are italicized.

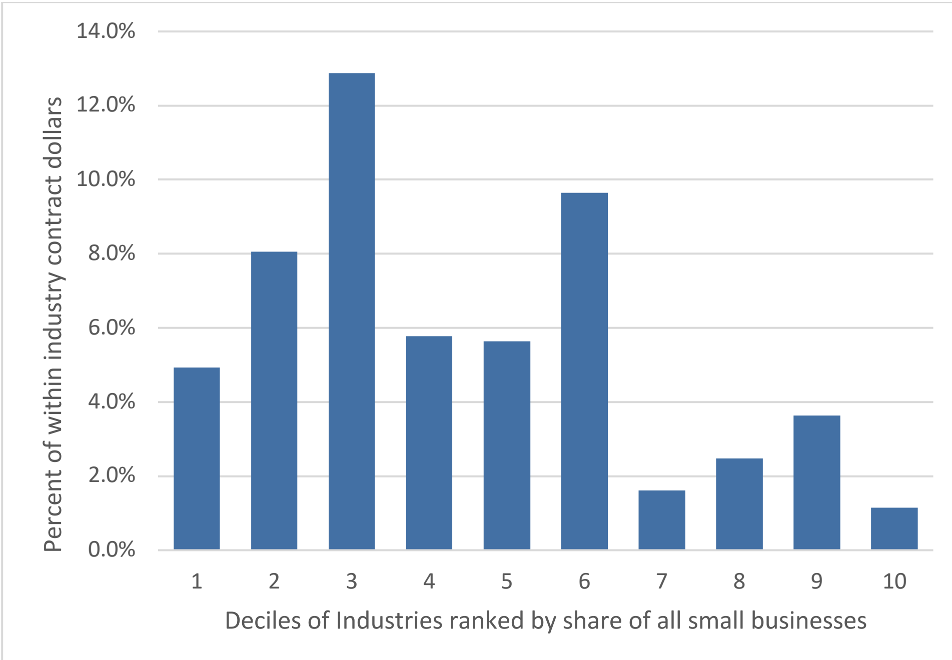


Figure C1. Alternate calculation of small Business Share of Within-Industry Contract Dollars, by Decile of Ranked Four-Digit Industries. The first decile has the most number of participating small businesses. Author's calculation using USAspending.gov (2013) data

Table C3. List of 50 Highest Growth Industries by 4-Digit North American Industrial Classification System (NAICS) Code, Haltiwanger et. al. (2014)

4-digit NAICS code	Description
2111	Oil and Gas Extraction
5179	Telecommunications Resellers
7111	Performing Arts Companies
5239	Other Financial Investment Activities
5259	Other Investment Pools and Funds
8113	Commercial and Industrial Machinery and Equipment (except Automotive and Electronic) Repair and Maintenance
5613	Employment Services
3312	Steel Product Manufacturing from Purchased Steel
2212	Natural Gas Distribution

4-digit NAICS code	Description
4871	Scenic and Sightseeing Transportation, Land
4245	Farm Product Raw Material Merchant Wholesalers
2382	Building Equipment Contractors
6114	Business Schools and Computer and Management Training
3311	Iron and Steel Mills and Ferroalloy Manufacturing
5416	Management, Scientific, and Technical Consulting Services
4812	Nonscheduled Air Transportation
5232	Securities and Commodity Exchanges
5611	Office Administrative Services
4239	Miscellaneous Durable Goods Merchant Wholesalers
5312	Offices of Real Estate Agents and Brokers
5415	Computer Systems Design and Related Services
2371	Utility System Construction
5231	Securities and Commodity Contracts Intermediation and Brokerage
2383	Building Finishing Contractors
6222	Psychiatric and Substance Abuse Hospitals
4235	Metal and Mineral (except Petroleum) Merchant Wholesalers
3369	Other Transportation Equipment Manufacturing
4247	Petroleum and Petroleum Products Merchant Wholesalers
5414	Specialized Design Services
2389	Other Specialty Trade Contractors
5629	Remediation and Other Waste Management Services
2381	Foundation, Structure, and Building Exterior Contractors
2372	Land Subdivision
2379	Other Heavy and Civil Engineering Construction
5179	Other Telecommunications
3362	Motor Vehicle Body and Trailer Manufacturing
3314	Nonferrous Metal (except Aluminum) Production and Processing
7121	Museums, Historical Sites, and Similar Institutions
3336	Engine, Turbine, and Power Transmission Equipment Manufacturing
2361	Residential Building Construction
2122	Metal Ore Mining
7115	Independent Artists, Writers, and Performers
5191	Other Information Services
2362	Nonresidential Building Construction
2131	Support Activities for Mining
3241	Petroleum and Coal Products Manufacturing
5191	Internet Publishing and Broadcasting and Web Search Portals
4861	Pipeline Transportation of Crude Oil
5313	Activities Related to Real Estate
5171	Wired Telecommunications Carriers

Sourcee. Haltiwanger et. al. (2014)

Table C4. Linear Regression including all variables of interest

	(1)	(2)
No. of NAICS	0.202***	0.208***
	(0.04)	(0.04)
top40 industry	0.034	0.029
	(0.07)	(0.07)
No. of agencies	0.402***	0.435***
	(0.08)	(0.10)
No. of subagency		-0.177
		(0.12)
No. of contracts	0.001***	0.001***
	(0.00)	(0.00)
DOD contract	0.625***	0.630***
	(0.07)	(0.07)
0-5 years	0.336***	0.338***
	(0.09)	(0.09)
6-10 years	-0.249	-0.248
	(0.13)	(0.13)
11-15 years	-0.319*	-0.320*
	(0.14)	(0.14)
16-20 years	-0.216*	-0.215*
	(0.11)	(0.10)
Fy13emp.	0.014	0.015
	(0.01)	(0.01)
Special status	0.004	0.005
	(0.06)	(0.06)
_cons	10.051***	10.187***
	(0.40)	(0.40)
N	3088.000	3088.000
vce	robust	robust

Table C5. Linear Regression Without Number of NAICS Codes variable to see effect on number of contracts variable

	est1 b/se
No. of agencies	0.647***
	(0.13)
No. of subagencies	0.039
	(0.15)
Special-status	0.043
	(0.06)
No. of contracts	0.001***
	(0.00)
DOD	0.716***
	(0.07)
0-5 years	0.387***
	(0.09)
6-10 years	-0.254
	(0.13)
11-15 years	-0.315*
	(0.15)
16-20 years	-0.229*
	(0.11)
21+ years (reference)	
Fy13 emp.	0.012
	(0.01)
top40 industry	0.014
	(0.07)
_cons	9.944***
	(0.41)
N	3088.000
vce	robust

REFERENCES

Artificial Lawyer. (2019, June 4). *France bans judge analytics, 5 years in prison for rule breakers*. https://www.artificiallawyer.com/2019/06/04/france-bans-judge-analytics-5-years-in-prison-for-rule-breakers/

Association of Administrative Law Judges vs. Colvin, 7th Cir. (2001). http://media.ca7.uscourts.gov/cgi-bin/rssExec.pl?Submit=Display&Path=Y2015/D01-23/C:14-1953:J:Posner:aut:T:fnOp:N:1489577:S:0

Best, J., & Tiede, L. B. (2015, January 22). *Vacancy in justice: Analyzing the impact of overburdened judges on sentencing decisions*. Available at SSRN. https://dx.doi.org/10.2139/ssrn.2417348

Chen, D. L., Moskowitz, T. J., & Shue, K. (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics, 131*(3), 1181-1242.

Choi, S. J., Gulati, M., & Posner, E. A. (2012). What do federal district judges want? An analysis of publications, citations, and reversals. *The Journal of Law, Economics, and Organization, 28*(3), 518–549. https://doi.org/10.1093/jleo/ewq020

Christiansen, E. A. (2018, September 6). *Supreme Court holds ALJ appointment procedure at SEC unconstitutional*. https://www.americanbar.org/groups/litigation/publications/litigation-news/top-stories/2018/sc-holds-alj-sec-appointment-procedure-unconstitutional/

Confidence in institutions: Gallup historical trends. (n.d.). http://news.gallup.com/poll/1597/confidence-institutions.aspx

Copus, R., & Hubert, R. (2017, October 17). *Detecting inconsistency in governance*. Available at SSRN. http://dx.doi.org/10.2139/ssrn.2812914

Cornell Law School's Legal Information Institute. (n.d.). *13 CFR § 121.102 - How does SBA establish size standards?* https://www.law.cornell.edu/cfr/text/13/121.102

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*(17), 6889-6892.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627–668.

Dirks, C. D. (2015, September 29). *Case law update: Federal circuit holds administrative law judges may be removed based on productivity statistics*. https://fedmanager.com/news/federal-circuit-holds-administrative-law-judges-may-be-removed-based-on-productivity-statistics?rq=ssa%20alj

D-U-N-S number. (n.d.) In Dun & Bradstreet, *Business glossary: Dun & Bradstreet D-U-N-S® number*. https://businesscredit.dnb.com/business-glossary/dun--bradstreet-duns-number/

eCFR. (n.d.). *Electronic Code of Federal Regulations: Title 48 Federal Acquisition Regulations System*. https://gov.ecfr.io/cgi-bin/text-idx?SID=11b42f278c5f115aec62cf72c472a3d9&mc=true&tpl=/ecfrbrowse/Title48/48tab_02.tpl

Epstein, L., Landes, W., & Posner, R. (2013). *The behavior of federal judges: A theoretical and empirical study of rational choice*. Cambridge, MA: Harvard University Press.

Farmer, L. (2019, July). *How state procurement offices are learning to buy better*. https://www.governing.com/topics/finance/gov-procurement-governing-institute-report.html

Federal Procurement Data System – Next Generation. (n.d.). *Top 100 contractors report*. https://www.fpds.gov/fpdsng_cms/index.php/en/reports/62-top-100-contractors-report

French, E., & Song, J. (2014). The effect of disability insurance receipt on labor supply. *American Economic Journal: Economic Policy, 6*(2), 291-337.

Friedman, M., & Friedman, R. D. (1962). *Capitalism and freedom*. University of Chicago Press.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics, 115*(3), 791-810. https://doi.org/10.1162/003355300554917

Gruhl, J., Spohn, C., & Welch, S. (1981). Women as policymakers: The case of trial judges. *American Journal of Political Science, 25*(2), 308-322. https://doi.org/10.2307/2110855

Haltiwanger, J. C., Hathaway, I., & Miranda, J. (2014). *Declining business dynamism in the U.S. high-technology sector* (SSRN Scholarly Paper No. ID 2397310). Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2397310

Haltiwanger, J. C., Jarmin, R. S., & Miranda, J. (2010). *Who creates jobs? Small vs. large vs. young* (Working Paper No. 16300). National Bureau of Economic Research. https://doi.org/10.3386/w16300

Haltiwanger, J. C., Jarmin, R. S., Kulick, R. B., & Miranda, J. (2016). *High growth young firms: Contribution to job, output and productivity growth* (U.S. Census Bureau Center for Economic Studies Paper No. CES-WP-16-49). Washington, DC: U.S. Census Bureau.

Heyes, A., & Saberian, S. (2019). Temperature and decisions: Evidence from 207,000 court cases. *American Economic Journal: Applied Economics, 11*(2), 238-265.

HQ/Parent D-U-N-S number. (n.d.). In Dun & Bradstreet, *Business glossary: HQ/Parent D-U-N-S® number*. https://businesscredit.dnb.com/business-glossary/hqparent-duns-number/

Hurst, E., & Pugsley, B. W. (2012). What do small businesses do? *Brookings Papers on Economic Activity*, *2011*(2), 73–118. https://doi.org/10.1353/eca.2011.0017

Hurst, E. G., & Pugsley, B. W. (2015). *Wealth, tastes, and entrepreneurial choice* (No. w21644). Cambridge, MA: National Bureau of Economic Research.

Jones, J. M. (2018, January 8). *Americans' identification as Independents back up in 2017*. https://news.gallup.com/poll/225056/americans-identification-independents-back-2017.aspx

Kueng, L., Yang, M. J., & Hong, B. (2014). *Sources of firm life-cycle dynamics: Differentiating size vs. age effects* (No. w20621). Cambridge, MA: National Bureau of Economic Research.

Lamm, D. V. (1988). Why firms refuse DOD business: An analysis of rationale. *National Contract Management Journal, 21*(2), 45.

Langbert, M., Quain, A. J., & Klein, D. B. (2016). Faculty voter registration in economics, history, journalism, law, and psychology. *Econ Journal Watch, 13*(3), 422-451.

Lucia et al. v. Securities and Exchange Commission, 585 U.S. 17-30 (2018). https://www.supremecourt.gov/opinions/17pdf/17-130_4f14.pdf

Maestas, N., Mullen, K. J., & Strand, A. (2013). Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. *American Economic Review, 103*(5), 1797-1829.

Marchand, R., & Warshawsky, M. (2015, September 9). *Reforming the system of review by administrative law judges in disability insurance* (Mercatus Working Paper). Available at SSRN. http://dx.doi.org/10.2139/ssrn.3191391

Marinescu, I. (2011). Are judges sensitive to economic conditions? Evidence from UK employment tribunals. *ILR Review*, *64*(4), 673–698. https://doi.org/10.1177/001979391106400403

Martin, J. D., & Sayrak, A. (2003). Corporate diversification and shareholder value: A survey of recent literature. *Journal of Corporate Finance*, *9*(1), 37-57.

Mashaw, J., Goetz, C., Goodman, F., Schwartz, W., Verkuil, P., & Carrow, M. (1978). *Social Security hearings and appeals: A study of the Social Security Administration hearing system*. Lexington, MA: Lexington Books.

Miller, M. A., & Ellis, J. M. (2016). *The NAICS code selection process and small business participation*. Monterey, CA: Naval Postgraduate School.

Nagle, J. F. (1999). *A history of government contracting*. Washington, DC: George Washington University Law School, Government Contracts Program.

Nakosteen, R., & Zimmer, M. (2014). Approval of social security disability appeals: Analysis of judges' decisions. *Applied Economics, 46*(23), 2783-2791. https://doi.org/10.1080/00036846.2014.914147

Newport, F. (2011, November 3). *Americans trust small-business owners most on job creation: President Obama's views generate more trust than those of congressional leaders*. http://news.gallup.com/poll/150545/americans-trust-small-business-owners-job-creation.aspx

Office of the Inspector General, Social Security Administration. (2017, September). *Audit report: Factors related to decreased administrative law judge productivity (A-12-*

18-50289). https://oig.ssa.gov/sites/default/files/audit/full/pdf/A-12-18-50289.pdf

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics, 11*(6), 619-632.

Public Works and Government Services Canada. (n.d.). *Standard acquisition clauses and conditions (SACC) manual: Chapter 1 - Public procurement*. https://buyandsell.gc.ca/policy-and-guidelines/supply-manual/section/1

Schrag, P., Schoenholtz, A. J., & Ramji-Nogales, J. (2009). *Refugee roulette: Disparities in asylum adjudication and proposals for reform*. New York, NY: NYU Press.

Shyu, J., & Chen, Y. L. (2009). Diversification, performance, and the corporate life cycle. Emerging Markets Finance and Trade, 45(6), 57-68.

Social Security Administration. (n.d.). Selected *data from Social Security's Disability Program: Applications for disability benefits and number awarded*. https://www.ssa.gov/oact/STATS/dib-g1.html

Social Security Administration. (2007, May 23). *Social Security testimony before Congress: Statement of Michael J. Astrue, Commissioner of Social Security: Testimony before the Senate Finance Committee.* https://www.ssa.gov/legislation/testimony_052307.html

Social Security Administration. (2012, June 27). *Social Security testimony before Congress: Statement of Michael J. Astrue, Commissioner, Social Security Administration before the Committee on Ways and Means Subcommittee on Social Security*. https://www.ssa.gov/legislation/testimony_062712.html

Social Security Administration. (2014, February 7). *HALLEX I-2-4-5: Dismissing a request for hearing*. https://www.ssa.gov/OP_Home/hallex/I-02/I-2-4-5.html

Social Security Administration. (2016, January 13). *CARES Plan: Leading the hearings and appeals process into the future: A plan for Compassionate And REsponsive Service*. https://www.ssa.gov/appeals/documents/cares_plan_2016.pdf

Social Security Administration. (2019, April 9). *HALLEX I-2-1-55: Assignment of service area cases to administrative law judges*. https://www.ssa.gov/OP_Home/hallex/I-02/I-2-1-55.html

Social Security Administration. (2020). *Hearings and appeals: ALJ disposition data FY 2020*. https://www.ssa.gov/appeals/DataSets/03_ALJ_Disposition_Data.html

Sunstein, C. R., Schkade, D., Ellman, L. M., & Sawicki, A. (2006). *Are judges political? An empirical analysis of the federal judiciary*. Washington, DC: Brookings Institution Press.

Wilson, J. Q. (1989). *Bureaucracy: What government agencies do and why they do it*. Basic Books.

U.S. Census Bureau. (n.d.). *North American Industry Classification System - Downloadable files*. https://www.census.gov/eos/www/naics/downloadables /downloadables.html

U.S. Census Bureau. (2015). *2012 SUSB annual datasets by establishment industry*. https://www.census.gov/data/datasets/2012/econ/susb/2012-susb.html

U.S. Census Bureau. (2020). *North American Industry Classification System (NAICS) main page: Introduction to NAICS*. https://www.census.gov/eos/www/naics/

U.S. Small Business Administration. (n.d.). *Market Research Guide for Contracting Officers* (Online course). https://www.sba.gov/course/market-research-guide-contracting-officers/#course-transcript

U.S. Small Business Administration. (2014, July 14). *Table of small business size standards matched to North American Industry Classification System Codes*. https://www.sba.gov/sites/default/files/Size_Standards_Table.pdf

U.S. Small Business Administration Office of Advocacy. (2016). *United States small business profile*. https://www.sba.gov/sites/default/files/advocacy/United_States.pdf

USAspending.gov. (2018). *Award data archive*. https://www.usaspending.gov/#/download_center/award_data_archive

BIOGRAPHY

Mary Clare Peate has a Master's in Education from Loyola Marymount University and a Master's in Public Policy from The University of Chicago.