The Journal of Private Enterprise 30(2), 2015, 83–102

The Effect of Tax Code Complexity on Entrepreneurship

Rick Weber*
Farmingdale State College

Abstract
The United States federal tax code is notoriously complex, and although there are many estimates of the costs of that complexity, there is little discussion of the effects of complexity on entrepreneurial activity. I use the word count of tax codes as a proxy for complexity to learn how complexity impacts entrepreneurship across the U.S. states. I find that a standard deviation increase in tax code length is associated with up to a 5 percent decline from the mean in business entry and exit rates. I conclude that tax code complexity has a negative impact on the dynamism of economies.

JEL Codes: K34, H32, L26, M13, O43
Keywords: tax code complexity, entrepreneurship, business exit, business entry

I. Introduction

The federal tax code in the United States is incredibly complex. However, little attention has been paid to complexity in states' tax codes. Gupta and Mills (2003) estimate that the 1,000 largest public firms face state tax compliance costs of around $300 million and federal tax compliance costs of around $1 billion. These compliance costs represent approximately 2.9 percent and 1.4 percent of their state and federal tax expenses, respectively (Gupta and Mills 2003). The cost of tax code complexity likely generates economies of scale due to expertise or experience in navigating required paperwork, so that the relative importance of these costs is likely higher for new ventures. Dealing with a complex tax system may be outsourced to an accounting firm, but it still adds a fixed cost that increases the minimum efficient scale of a business.

* I thank participants at the Free Market Institute research workshop, the Suffolk University economics department workshop, the Association of Private Enterprise Education conference, and the Public Choice Society conference for helpful comments. Additionally, I would like to thank Benjamin Powell for helpful comments on early drafts of this paper.

What is not clear is how large an impact such complexity has on entrepreneurial activity. Complex tax codes affect American entrepreneurs from the federal level down to the local level and add compliance costs at each step. This study will examine the effect of state tax code complexity. Data collected for this research suggest that such complexity is a fraction of that at the federal level, with the Internal Revenue Code estimated to be over 3.4 million words (Walker 2013). Because this study allows a cross-sectional comparison, it will offer insight into the economic effect of tax code complexity. The findings of this study suggest that tax code complexity at the state level matters, and by extension that federal tax code complexity may have a profound impact on entrepreneurs.

There are at least three big questions to ask about the complexity of a tax code: (1) How do we measure complexity? (2) What is the impact of complexity on economic agents? (3) What causes this complexity? This paper investigates the first two questions. I use word count as a proxy for complexity and regress business entry rates and business exit rates on word count (as well as control variables).

There is an extensive literature investigating entrepreneurship, institutional quality, and economic growth, but the literature on tax code complexity is relatively small and there are few papers connecting the two. This paper fills that gap with a comparative study of tax code complexity across states and the impact of that complexity on levels of entrepreneurship. This paper also introduces an innovation in the empirical entrepreneurship literature by examining business failures rather than just startup activity. By ignoring failure, the existing literature misses an important variable that is an inescapable part of a dynamic economy.

There are three common technical conceptions of entrepreneurship (Bjørnskov and Foss 2008). In the Schumpeterian view, entrepreneurs innovate, and in doing so, they breed creative destruction (Schumpeter 1911). The Kirznerian view is that they can discover profit opportunities, and in their alertness drive the process of market equilibration (Kirzner 1978). The Knightian view is that they make judgments in the face of uncertainty (Knight 1921). All three conceptions are complementary and capture important facets of entrepreneurship, but the Schumpeterian view is especially important for this study. Specifically, business failure captures the "destruction" in creative destruction and it is this aspect of Schumpeter's ideas to which the title of this paper alludes. Baumol (1990) points out the need to distinguish between productive and destructive

entrepreneurship; the former creates wealth through mutually advantageous exchanges while the latter makes a few entrepreneurs rich by imposing costs on others. Lobbying for tax loopholes is an example of destructive entrepreneurship. Understanding this possibility is important for interpreting the results of this study.

The theory of entrepreneurship is well developed, but measuring entrepreneurship is a difficult proposition. Because there are many aspects of entrepreneurship, there are a number of places to look for entrepreneurship. I use a measure of business entry rates as used in Campbell and Rogers (2007), Bacher and Brülhart (2013), Sobel (2008), and Powell and Weber (2013). Unlike the cited studies, I supplement this measure with business exit rates to get a better view on the dynamism of Schumpeterian entrepreneurship. Existing literature measures entrepreneurial activity by looking at positive measures of entrepreneurial activity (such as business entry), but using entry alone paints an incomplete picture. Looking at failure is important because it gets at the destruction in Schumpeter's idea of creative destruction. In 1900, the United States had over 10,000 carriage manufacturers (Carriage Museum of America 2013). These businesses were displaced by a smaller number of automobile manufacturers; an estimated 1,500 automobile manufacturers have operated in the United States since 1896 (Dreyer 2009). This case makes it clear that a business unable to compete with entrepreneurial newcomers may fail, and that the number of failed firms may outnumber the number of entrants. While other factors may cause business failure, the appropriate statistical controls can allow us to safely conclude that failure is indicative of entrepreneurship rather than, for example, a general economic slump, or retirement by a firm's owner-operator. Macroeconomic and demographic controls used in this study help to control for such issues. With these caveats in mind, business failure can be thought of as a symptom of entrepreneurial success (of other businesses) that frees up resources for innovators to expand on their success. This logic is confirmed by Sobel, Clark, and Lee (2007), who find that economic freedom is positively associated with entrepreneurship as measured by the Global Entrepreneurship Monitor and business failure rates.

As economists, we are concerned with entrepreneurship because it is the driving force of the market process. Entrepreneurship is fundamentally about experimentation and discovery, and failure is an integral part of this overall process. In seeking profit opportunities, entrepreneurs arbitrage price differences, make use of innovations,

and reallocate resources to more accurately reflect the desires of consumers and the opportunity costs of productive factors. Companies' shrinking and ultimately going out of business is a part of the process of moving factors of production to their highest valued use. If inefficient businesses are not shutting down in the face of "the perennial gale of creative destruction," this is a sign that something is impeding the entrepreneurial process.

Yamakawa, Peng, and Deeds (2010) offer an important contribution in which they find that, under certain conditions, individual entrepreneurs learn from their failed enterprises. This learning allows them to be more successful in subsequent ventures, showing that entrepreneurial failure is also an input to future success. In addition to the beneficial aspects of failure discussed above, we can think of business failure as a source of hands-on education ("the school of hard knocks").

There has been much research on the connection between institutional quality and entrepreneurship.[1] Studies in this literature that look at the United States typically use the Economic Freedom of North America index to measure institutional quality (Avilia, Ashby, and McMahon 2012). It includes measures of the size of government, taxation, and labor market freedom. Kreft and Sobel (2005) and Sobel (2008) find a strong correlation between EFNA scores and entrepreneurship, as well as between entrepreneurship and living standards.

But these measures of institutional quality and economic freedom do not look at tax code complexity. There is a literature examining the impact of tax code complexity on costs of compliance (e.g., Gupta and Mills 2003) and levels of compliance (e.g., Forest and Sheffrin 2002). These measures are important, but overlook a third facet of this issue. When you see a complicated tax code, you have three options: deal with it (in which case we want to know about compliance costs), cheat (in which case we are discussing compliance levels), or walk away entirely. This third issue is the focus of this paper. How does complexity affect entrepreneurs' decision to go into (or remain in) business? This issue has implications for the overall dynamism of a market beyond issues of static economic efficiency and public finance.

[1] For example, see Bjørnskov and Foss (2008, 2012), Campbell and Rogers (2007), Hall and Sobel (2008), Kreft and Sobel (2005), Powell and Weber (2013), Sobel, Clark, and Lee (2007), and Sobel (2008).

Bacher and Brülhart (2013) investigate the effects of Swiss canton and municipality tax codes including the effect of tax code complexity. They use the number of tax brackets and the number of words in each code to measure complexity and find that more words are associated with lower levels of entrepreneurship (as measured by firm entry). Edmiston, Mudd, and Valev (2003) examine the effects of complexity on foreign direct investment in transition economies, using the number of lines discussing the tax base (in the *Central and East European Tax Directory*), the number of different tax rates, and an index of the presence of indefinite language to measure complexity. They find complexity to have a negative effect on foreign direct investment.

This paper compares complexity across states and is the first paper, to my knowledge, that uses the state tax codes as a data source. By examining the connection between tax codes and entrepreneurship, I contribute a first step toward using information on tax code complexity to refine our understanding of economic institutional quality. This paper also provides a unique contribution by including business failure as well as business starts to get a richer view of entrepreneurship that is more in line with Schumpeter's theoretical conception of entrepreneurship as an innovative and disruptive force.

II. Data and Methodology

I estimate the effect of tax code complexity on entrepreneurship by using business entry and exit rates as dependent variables. These data are from the Census Bureau's *Business Dynamics Statistics* project, which includes information based on firm or establishment characteristics. In this context, a firm is an overarching business unit (e.g., McDonald's) that is comprised of and controls one or more establishments (e.g., the McDonalds on 19th street). Each establishment is a physical location that undertakes business operations. For this study, I used the dataset based on firms, which shows (for example) the entry of establishments under the umbrella of a firm of a certain size. This allows me to draw conclusions about changes in business entry and exit with an eye toward a larger firm's ability to support to its establishments. Ideally, this study would account for firm characteristics such as age and industry. Unfortunately, the data do not allow for such detail in a cross-section of the states.

The main explanatory variable of interest is tax code complexity, which is not straightforward to define. Many factors affect the complexity of a tax code, such as loopholes, the breadth of the tax base, and the language of the code. In general, complexity is the combination of traits that raise difficulties in calculating one's tax liability.

Complexity may manifest itself in a number of ways. A large tax code—one with many words or sections—is complex. Vague language is an important source of complexity as it creates uncertainty about the outcomes of certain actions. A more specific code would also be larger, but the results of this study imply that longer tax codes in this dataset are more complex and that the negative effect of this complexity outweighs any positive effect of clarity. I use the number of words in a state's tax code as a proxy for complexity. To calculate word counts, I compiled each state's tax code from the HTML version available online into a single text file. Most states, when posting their tax codes online, split the document up by sections, chapters, and articles so that the entire tax code may be split up over thousands of individual web pages. Consequently, I only included states for which I could automate the downloading process, resulting in a sample of thirty-three states.

Table 1. Comparison of States in Sample with Missing States

Variable	In sample mean (n=33)	Out of sample mean (n=17)	Overall standard deviation
Tax burden	3.38	3.52	0.98
Private industry GSP per capita	$43,025	$43,644	$15,732
Unemployment	9.18%	9.68%	2.16%
% of pop. with bachelor's degree	28.65%	26.31%	4.84%
Median age	37.39	37.94	2.27
% of population that is white	77.54%	75.19%	12.94
Population density	196.58	191.81%	261.09
Establishment entry rate	9.80%	9.31%	1.26%
Establishment exit rate	6.45%	6.38%	1.06%

Table 1 presents unweighted mean values for important variables used in this study for the sample states and the excluded states. The means are all within one standard deviation of each other and a differences-in-means test confirms that the mean values are not statistically significantly different. Although many excluded states are in the South, the sample appears to represent the country as a whole.

Table 2 lists the word count of included states' tax codes. The mean code is approximately 450,000 words, and the standard deviation is approximately 250,000. Most states' codes are clustered around 300,000 words, though another group is clustered around 550,000 words. California and New York both have codes longer than one million words; when they are removed from the sample, the mean is 405,000 words, with a standard deviation of 175,000 words. At an estimated 3.4 million words, the federal tax code is an order of magnitude larger than the average state's code.

For this study, I modeled the effect of tax code complexity as a linear relationship between the number of words in a state's tax code (wc_i) and the level of entrepreneurship in that state. I used a cross section with all variables from the year 2010, except for tax code word counts, which were calculated with the most recent data available.

$$Ent_i = \alpha + \beta_0 wc_i + \beta_1 controls_i + \varepsilon$$

Figure 1 data model

It is important to bear in mind that tax code complexity may take the form of loopholes that reduce tax burden (as argued by Bruce and Gurley-Calvez [2006]). As such, it is important to control for tax burden. Rather than using tax rates, I use area 2 from the Economic Freedom of North America (EFNA) index (using the state and local measures rather than the measures that include the effects of the federal government). This measure compares states on takings and discriminatory taxation which includes measures of total tax revenue and indirect tax revenue as a percentage of gross state product (GSP), as well as top marginal income tax rates. The EFNA index and its subcomponents are reported as a score between 0 and 10. To ease interpretation, I subtract each area 2 score from ten (the maximum possible score) so that an increase in the variable can be interpreted as an increased tax burden. Essentially, I use an aggregated measure of average tax rates rather than specific measures of marginal tax rates. In addition to being easier, it makes more sense to use average taxes with business entry rates since this measure of entrepreneurship focuses on the decision to go into (or out of) business rather than the decision of whether to do more of an activity already engaged in (i.e., the decision to expand one's existing business).

Table 2. States' Tax Code Lengths

State	Word Count
Alaska	234,549
Arizona	209,843
California	1,275,607
Connecticut	512,090
Delaware	172,054
Florida	585,533
Hawaii	272,189
Idaho	204,850
Illinois	506,866
Indiana	746,415
Kansas	381,835
Kentucky	355,579
Louisiana	825,829
Massachusetts	363,577
Michigan	563,616
Minnesota	557,013
Missouri	309,817
Montana	263,403
Nebraska	360,985
Nevada	298,050
New Hampshire	151,890
New York	1,018,035
North Dakota	303,440
Ohio	666,277
Oregon	514,322
Rhode Island	369,929
South Dakota	232,989
Texas	483,029
Utah	374,822
Vermont	252,597
Virginia	643,480
Washington	329,586
Wisconsin	528,243

Source: Individual state tax code websites linked to from FindLaw.com.

In addition to these variables, I include economic and demographic control variables that are standard in the entrepreneurship literature. These are civilian unemployment rates, median age, percentage of the population that is white, percentage of the population with a bachelor's degree, population density, GSP per capita, and civilian unemployment rate.

Older populations and populations with a higher presence of racial minorities are typically less entrepreneurial (see, for example, Campbell and Rogers [2007] and Sobel [2008]) while higher levels of postsecondary schooling are associated with lower levels of entrepreneurship (Kreft and Sobel [2005]). Unemployment rates and GSP per capita have varying relationships with levels of entrepreneurship. Including them is important to control for macroeconomic conditions, but their correlations with entrepreneurship say more about those conditions than the underlying relationship between the variables. The summary statistics for word count, entry rates, and exit rates is included in table 3. Summary statistics for the remaining variables are available upon request from the author.

Table 3. Select Summary Statistics

Variable	Source	Mean	Standard Deviation
Word Count	Compiled from individual tax codes, from "State Tax Codes," FindLaw.com	450,556	248,830
Establishment Entry Rate—Firm Data (2010)	Business Dynamics Statistics, United States Census Bureau		
	All firms	9.80%	1.37%
	Small (fewer than 50 employees)	10.91%	1.69%
	Medium (50–249 employees)	3.91%	1.27%
	Large (250 or more employees)	7.02%	0.99%
Establishment Exit Rate—Firm Data (2010)	Business Dynamics Statistics, United States Census Bureau		
	All firms	10.34%	1.54%
	Small (fewer than 50 employees)	12.35%	1.96%
	Medium (50–249 employees)	1.96%	0.60%
	Large (250 or more employees)	4.89%	0.94%

III. Results: A First Look

All regressions in this paper are presented with standardized coefficients representing the impact of increasing a variable by one standard deviation in terms of standard deviation changes in the

dependent variable. This presentation eases comparison of the relative importance of different variables.

Table 4 presents the main regression results. Because California and New York both have substantially longer tax codes than the rest of the sample, I also ran the regressions excluding them as outliers. These states also had high entry and exit rates, averaging 10.9 percent and 11.3 percent, respectively, compared with 9.8 percent and 10.3 percent for the remaining states. These are large state economies that appear to diverge from the effects seen in the rest of the sample. Further research is necessary to understand how these states differ from the rest and the causes of such differences.

The coefficients for both entry and exit rates are negative, though statistically insignificant at the traditional thresholds when the full sample is used (although they are significant at the 35 percent and 15 percent levels for entry and exit, respectively). The adjusted R^2 for the entry and exit rate regressions were 0.12 and 0.47, respectively. The F-statistic for the entry rate regression is insignificant, but significant at the 1 percent level for exit rates. Excluding California and New York increases the significance of the coefficients for both entry and exit rates. The adjusted R^2 for entry and exit rates rises to 0.46 and 0.67, respectively. The significance of the F-statistics also rises, with the entry rate regression becoming significant at the 10 percent level. The word count coefficients are both approximately −0.36 and are significant at the 5 percent level. These coefficients imply that a standard deviation increase in word count (approximately 175,000 words when the outliers are excluded) is associated with a fall in entry and exit rates of 0.49 percent and 0.55 percent, respectively (or approximately 5 percent of the mean entry and exit rates). For the mean state, this means 559 fewer new establishments entering the market and 636 establishments *not* going under that might otherwise have been expected to.

For example, Texas has a tax code of 483,000 words, but if it were 87,500 words longer (approximately one-half of a standard deviation), its predicted annual establishment entry level would fall by 2,320 establishments. Its predicted annual establishment exit level, which represents businesses likely to be weeded out by the competitive market process, would fall by 2,604 establishments.

Table 4. Main Results, All Firms

All Firms	**Full Sample**		**Excluding NY & CA**	
Dependent variable				
(standard errors in parentheses)	Entry rate	Exit rate	Entry rate	Exit rate
word count	−0.1951	−0.2373	−0.3556 **	−0.3602 **
	(0.2034)	(0.1585)	(0.1680)	(0.1302)
tax burden	−0.0516	−0.01436	−0.2827	−0.2068
	(0.2733)	(0.2130)	(0.2594)	(0.2011)
unemployment	0.4475 **	0.7187 ***	0.4298 **	0.6958 ***
	(0.2074)	(0.1616)	(0.1913)	(0.1483)
bachelor's degree	0.1572	0.2809	0.1264	0.2487
	(0.2325)	(0.1812)	(0.2182)	(0.1692)
median age	−0.4136 **	−0.3636 **	−0.3684 *	−0.3205 **
	(0.1985)	(0.1547)	(0.1882)	(0.1459)
percent white	−0.04991	−0.1609	−0.06348	−0.1647
	(0.2055)	(0.1602)	(0.1851)	(0.1435)
GSP/capita	0.01112	−0.2874	−0.2191	−0.4738 **
	(0.2907)	(0.2265)	(0.2845)	(0.2205)
pop. density	−0.1886	−0.1008	−0.07306	0.002135
	(0.2397)	(0.1868)	(0.2311)	(0.1791)
Observations	33	33	31	31
R2	0.3403	0.5994	0.4584	0.6745
Adjusted R2	0.1204	0.4659	0.2614	0.5562
F statistic	1.548 (8,24)	4.489 (8,24)	2.327 (8,22)	5.699 (8,22)
F stat p-value	0.8068	0.9980	0.9441	0.9995
Significance	*10%	** 5%	*** 1%	

Among the control variables, unemployment rates and median age are consistently statistically significant. Median age is consistently negative, with coefficients ranging between −0.32 and −0.42. These results imply that increasing a state's median age by 2.4 years will reduce its predicted entry rate by approximately 0.6 percent and its exit rate by approximately 0.5 percent. Unemployment yields a positive coefficient for all regressions (at the 5 percent level for entry and 1 percent for exit), implying that a standard deviation increase in unemployment rate (2.2 percent for the full sample) is associated with

an increase in establishment entry rates of approximately 0.6 percent, and an increase in exit rates of approximately 1 percent. As stated in the data section, it is difficult to draw conclusions about causation from these results. What we can say is that in 2010, higher unemployment rates were associated with significant business turnover. Per capita GSP is significant in the exit rate regression for the restricted sample and implies that an increase in per capita GSP of $7,960 will reduce predicted exit rates by approximately 0.7 percent.

IV. Results by Firm Size

I also broke the data down by firm size and ran regressions on three groups based on the number of employees in the overarching firm. Tables 5, 6, and 7 present the complexity coefficients. Small firms (table 5) are defined here as having fewer than 50 employees, medium (table 6) as having 50 to 249, and large (table 7) as having 250 or more. The results for medium-sized firms are statistically insignificant, but only cover approximately 6 percent of the establishments in the sample. By contrast, establishments within large firms represent nearly 19 percent of all establishments in the sample, and small firms represent the remaining 75 percent.

The word count coefficients for large and small firms remain negative across the board. The coefficients for small firms are roughly the same magnitude as those reported earlier, which is to be expected since small firms make up such a large portion of the sample. The results with the full sample remain insignificant, although the adjusted R^2 and F-statistics indicate a better fit than the "all firms" regressions, as can be seen in the tables. The word count coefficients when outliers are excluded are −0.33 for entry rates (significant at the 10 percent level) and −0.32 for exit rates (significant at the 5 percent level). The adjusted R^2 is also higher than the "all firms" regressions, increasing to 0.50 and 0.68 for entry and exit rates, respectively. F-statistics increase in significance to the 5 percent level for entry rates and remain significant at the 1 percent level for exit rates. Statistical significance and magnitude of control variables are approximately the same for the small firms and all firms regressions.

Table 5. Results for Firms with Fewer Than 50 Employees

	Full Sample		**Excluding NY & CA**	
Dependent variable				
(standard errors in parentheses)	Entry rate	Exit rate	Entry rate	Exit rate
word count	−0.1877	−0.2542	−0.3292 *	−0.3184 **
	(0.1945)	(0.1501)	(0.1611)	(0.1288)
tax burden	−0.06173	−0.1084	−0.2723	−0.2429
	(0.2614)	(0.2018)	(0.2488)	(0.1989)
unemployment	0.4697 **	0.7187 ***	0.4512 **	0.6911 ***
	(0.1983)	(0.1531)	(0.1835)	(0.1467)
bachelor's degree	0.1091	0.2	0.0815	0.1731
	(0.2224)	(0.1716)	(0.2092)	(0.1673)
median age	−0.4339 **	−0.3353 **	−0.3915 **	−0.2997 **
	(0.1898)	(0.1465)	(0.1804)	(0.1443)
percent white	−0.02531	−0.2252	−0.03889	−0.222
	(0.1966)	(0.1517)	(0.1775)	(0.1419)
GSP/capita	0.0115	−0.3105	−0.2001	−0.4421 *
	(0.2781)	(0.2146)	(0.2728)	(0.2181)
pop. density	−0.2099	−0.04968	−0.1043	0.02708
	(0.2292)	(0.1769)	(0.2216)	(0.1772)
Observations	33	33	31	31
R2	0.3965	0.6405	0.5020	0.6816
Adjusted R2	0.1954	0.5206	0.3209	0.5658
F statistic	1.971 (8,24)	5.344 (8,24)	2.772 (8,22)	5.886 (8,22)
F stat p-value	0.9050	0.9994	0.9723	0.9996
Significance	*10%	** 5%	*** 1%	

Table 6. Results for Firms with 20 to 249 Employees

	Full Sample				Excluding NY & CA			
Dependent variable	Entry rate		Exit rate		Entry rate		Exit rate	
(standard errors in parentheses)								
word count	0.1593		0.05766		0.03407		0.1039	
	(0.1743)		(0.1754)		(0.1532)		(0.1595)	
tax burden	−0.4636	*	−0.2837		−0.5393	**	−0.2126	
	(0.2342)		(0.2358)		(0.2367)		(0.2464)	
unemployment	−0.09629		0.552	***	−0.08605		0.5491	***
	(0.1777)		(0.1789)		(0.1745)		(0.1817)	
bachelor's degree	−0.002482		0.1489		−0.001035		0.1642	
	(0.1993)		(0.2006)		(0.1991)		(0.2072)	
median age	−0.1004		−0.07368		−0.09217		−0.09299	
	(0.1701)		(0.1712)		(0.1716)		(0.1787)	
percent white	−0.6604	***	−0.2209		−0.648	***	−0.2129	
	(0.1761)		(0.1773)		(0.1688)		(0.1758)	
GSP/capita	−0.4128		−0.3893		−0.5167	*	−0.3352	
	(0.2492)		(0.2508)		(0.2595)		(0.2701)	
pop. density	−0.04357		0.2338		−0.004501		0.1972	
	(0.2054)		(0.2068)		(0.2108)		(0.2194)	
Observations	33		33		31		31	
R2	0.5154		0.5090		0.5494		0.5116	
Adjusted R2	0.3539		0.3454		0.3855		0.3340	
F statistic	3.191 (8,24)		3.11 (8,24)		3.352 (8,22)		2.881 (8,22)	
F stat p-value	0.9870		0.9852		0.9884		0.9766	
Significance	*10%		** 5%		*** 1%			

Table 7. Results for Firms with 250 or More Employees

	Full Sample		Excluding NY & CA	
Dependent variable				
(standard errors in parentheses)	Entry rate	Exit rate	Entry rate	Exit rate
word count	−0.4084 *	−0.3104	−0.2964	−0.3555 **
	(0.2061)	(0.1902)	(0.1871)	(0.1674)
tax burden	0.09366	−0.1148	0.09495	−0.2445
	(0.2770)	(0.2555)	(0.2889)	(0.2586)
unemployment	0.3493	0.325	0.3247	0.3014
	(0.2102)	(0.1939)	(0.2131)	(0.1907)
bachelor's degree	0.2096	0.1706	0.1907	0.1412
	(0.2356)	(0.2174)	(0.2430)	(0.2175)
median age	−0.09356	0.03041	−0.0792	0.06696
	(0.2011)	(0.1856)	(0.2096)	(0.1875)
percent white	−0.05978	−0.4148 **	−0.04758	−0.401 **
	(0.2083)	(0.1922)	(0.2062)	(0.1845)
GSP/capita	0.2301	0.1381	0.2414	0.005975
	(0.2946)	(0.2718)	(0.3168)	(0.2835)
pop. density	0.1494	0.1078	0.1587	0.1841
	(0.2429)	(0.2241)	(0.2574)	(0.2303)
Observations	33	33	31	31
R2	0.3223	0.4232	0.3282	0.4619
Adjusted R2	0.0965	0.2310	0.0839	0.2663
F statistic	1.427 (8,24)	2.201 (8,24)	1.343 (8,22)	2.361 (8,22)
F stat p-value	0.7640	0.9354	0.7252	0.9471
Significance	*10%	** 5%	*** 1%	

Although restricting analysis to large firms does not change the sign of the complexity coefficients, it does affect the statistical significance and magnitude. In the full sample, the coefficient for establishment entry rates is significant at the 10 percent level with a standardized coefficient of −0.41, implying approximately 110 fewer new large establishments for the average state. In the restricted sample, the coefficient for exit rates is −0.36 and significant at the 5 percent level, implying 75 fewer large-establishment exits for the average state. However, the regressions for large firms perform

relatively weakly with adjusted R^2 under 0.1 for entry rates and below 0.3 for exit rates. F-statistics are significant at the 10 percent level for exit rates, but statistically insignificant for entry rates. Among the control variables, only the percent white coefficients are statistically significant, and only in the exit rate regressions. Both coefficients are approximately −0.4 and significant at the 5 percent level. This finding implies that increasing the population that is white by 13 percent will increase the predicted exit rate by approximately 0.3 percent.

For medium firms, word count yields relatively small, positive, statistically insignificant coefficients. However, the F-statistics are all significant at the 5 percent level. R^2 is approximately 0.35 for all four regressions. Tax burden is negatively and significantly associated with lower entry rates (as expected), higher unemployment is positively associated with exit rates (indicating poor macroeconomic conditions were hard on medium-sized firms in 2010), the percent white is negatively associated with entry rates (contrary to what is usually found in the literature), and per capita GSP is negatively associated with entry rates in the restricted sample. Tax burden coefficients are negative, as predicted, but they are not statistically significant, as I would have predicted.

V. Robustness Checks

In addition to the baseline regressions, I ran a variety of alternative specifications, and the findings were generally robust for the restricted sample. Increased tax code complexity is associated with lower levels of business failure and at least some reduction in establishment entry. However, California and New York warrant deeper analysis.

To see if tax code length might have a nonlinear relationship, I also altered the model to use the natural logarithm of word count and word count squared. When the full sample is used, the logarithmic model outperforms the quadratic model with adjusted R^2 of 0.14 and 0.51 (for entry and exit rates, respectively) for the logarithmic model and 0.097 and 0.43 for the quadratic model. In the logarithmic model, word count yields a negative and significant coefficient for exit rates, in line with the linear estimates and in spite of the inclusion of the outliers. Excluding New York and California again increased statistical significance, yielding significant word count coefficients for both models at similar magnitudes to the linear model. With the restricted sample, both models yield essentially the same results as the linear model in terms of magnitude of coefficients, statistical

significance, and overall model performance. Among the restricted sample states, it appears that a linear model is sufficient to capture the relationship between tax code complexity and entrepreneurship. Using the full sample, only the logarithmic model yields a statistically significant result, which lends credence to the view that tax code complexity has a particular impact on business exit rates.

I also ran regressions using the Shannon index of industrial specialization (Nissan and Carter 2010). This index measures the degree to which employment is concentrated in a small number of industries or spread more evenly across a larger number of industries. When included, it generally yielded coefficients that implied that lower levels of industrial concentration are associated with lower levels of business failures, but it did not affect the complexity coefficients substantively.

Eliminating variables with little explanatory power yields no substantive changes in the results. I ran regressions, including the percentage of a state's electorate voting for Mitt Romney in the 2012 presidential election, to control for partisan ideology. The variable was insignificant and did not appreciably affect the main results.

Under a variety of specifications, there was consistent evidence for longer tax codes being associated with lower levels of business failure. Under many specifications, there was statistically significant evidence of a fall in business entry, with negative coefficients remaining in specifications lacking statistical significance.

VI. Conclusion

This paper has been a first exploration of state tax codes' impact on entrepreneurship. Despite the data's limitations, I have found evidence that longer state tax codes are associated with lower establishment entry and exit rates. Caution is necessary in interpreting these results, as these effects are diminished with the inclusion of California and New York. Further research is necessary to understand the relative entrepreneurial success of these two economies.

Although the fall in exit rates seems like a reason to celebrate, the commensurate fall in entry rates implies that these states have less dynamic economies. The presence of business failure indicates healthy competition and turnover in the market. Business exits matter because they're part of a wider experimental and allocative *process*. One entrepreneur's success can result in the failure of less effective competitors. This process is important to free up factors of

production to best meet consumers' demands given scarcity. It is also important for introducing innovations to the market. As Schumpeter (1942) put it:

> In capitalist reality as distinguished from its textbook picture, it is not [price] competition which counts but the competition from the new commodity, the new technology, the new source of supply, the new type of organization . . . competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits and the outputs of the existing firms but at their foundations and their very lives.

Tax codes are part of a political-economic environment that shapes the incentives facing entrepreneurs. Long and arcane tax codes can increase would-be entrepreneurs' perception of uncertainty, which is consistent with a fall in entry rates. Complex codes also create opportunities for rent-seeking that might protect incumbent firms against competition, which is consistent with lower exit rates. Because of these possibilities, better understanding the causes and effects of tax code complexity is important for understanding the political institutions that shape the context entrepreneurs face. Determining if these hypotheses explain the evidence will require in-depth analysis of the political causes of tax code complexity and the effects of complexity on individual entrepreneurs.

A few caveats are necessary to bear in mind. These econometric measures are valid on the margin, but it is unlikely that the impact of tax code complexity is marginal. A high-complexity state will probably not increase its entry rate by shedding 1,000 words from its tax code; the change must either be big enough to be noticed or must come through other effects.

This study only looked at the state level, where the average tax code is 450,000 words, but entrepreneurs must also contend with the byzantine national code, which is millions of words long. If the national code has effects similar to those hypothesized here, it could have a profound negative effect on entrepreneurship in all states.

With these caveats in mind, it is clear that the evidence of this study understates the importance of tax code complexity. Further research must estimate the impact of the national tax code, as well as the impact of complexity on the composition of markets. Examining

tax code complexity internationally has difficulties (such as comparing tax codes while accounting for language differences and a lack of the shared institutional framework that the states have), but is important to get closer to finding the full impact of tax code complexity. Research on tax code complexity and destructive entrepreneurship would also be beneficial.

References

Bacher, Hans Ulrich, and Marius Brülhart. 2013. "Progressive Taxes and Firm Births." *International Tax and Public Finance*, 20(1): 129–68.

Baumol, William J. 1990. "Entrepreneurship: Productive, Unproductive, and Destructive." *Journal of Political Economy*, 98(5, pt. 1): 893–921.

Bjørnskov, Christian, and Nicolai J. Foss. 2008. "Economic Freedom and Entrepreneurial Activity: Some Cross-Country Evidence." *Public Choice*, 134(3–4): 307–28.

Bjørnskov, Christian, and Nicolai J. Foss. 2012. "How Institutions of Liberty Promote Entrepreneurship and Growth." *Economic Freedom of the World: 2012 Annual Report*. Vancouver, Canada: Fraser Institute.

Bruce, Donald, and Tami Gurley-Calvez. 2006. "Federal Tax Policy and Small Business." SSRN Scholarly Paper ID 904641. Rochester, NY: Social Science Research Network.

Bueno, Avilia, Nathan J. Ashby, and Fred McMahon. 2012. *Economic Freedom of North America: 2012*. Vancouver: Fraser Institute.

Campbell, Noel D., and Tammy M. Rogers. 2007. "Economic Freedom and Net Business Formation." *Cato Journal*, 27(1): 23–36.

Carriage Museum of America. 2013. "Carriage Manufacturers." Lexington, KY: Carriage Museum of America. Accessed November 13.

Dreyer, Rachael. 2009. "Automotive History." Ann Arbor, MI: Bentley Historical Library at the University of Michigan.

Edmiston, Kelly, Shannon Mudd, and Neven Valev. 2003. "Tax Structures and FDI: The Deterrent Effects of Complexity and Uncertainty." *Fiscal Studies*, 24(3): 341–59.

Forest, Adam, and Steven M. Sheffrin. 2002. "Complexity and Compliance: An Empirical Investigation." *National Tax Journal*, 55(1)): 75–88.

Gupta, Sanjay, and Lillian F. Mills. 2003. "Does Disconformity in State Corporate Income Tax Systems Affect Compliance Cost Burdens?" *National Tax Journal*, 56(2): 355–71.

Hall, Joshua C., and Russell S. Sobel. 2008. "Institutions, Entrepreneurship, and Regional Differences in Economic Growth." *Southern Journal of Entrepreneurship*, 1(1): 69–96.

Kirzner, Israel M. 1978. *Competition and Entrepreneurship*. Chicago: University of Chicago Press.

Knight, Frank H. 1921. *Risk, Uncertainty and Profit*. Boston: Houghton Mifflin.

Kreft, Steven F., and Russell S. Sobel. 2005. "Public Policy, Entrepreneurship, and Economic Freedom." *Cato Journal*, 25(3): 595–616.

Nissan, Edward, and George Carter. 2010. "Shannon Measure of Industrial Specialization at the State Level." *Southwestern Economic Review*, 37(1): 29–45.

Powell, Benjamin, and Rick Weber. 2013. "Economic Freedom and Entrepreneurship: A Panel Study of The United States." *American Journal of Entrepreneurship*, 6(1): 67–87.

Schumpeter, Joseph A. 1911[1934]. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Translated by Redvers Opie. Cambridge, MA: Harvard University Press.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper Collins.

Sobel, Russell S. 2008. "Testing Baumol: Institutional Quality and the Productivity of Entrepreneurship." *Journal of Business Venturing*, 23(6): 641–55.

Sobel, Russell S., J. R. Clark, and Dwight R. Lee. 2007. "Freedom, Barriers to Entry, Entrepreneurship, and Economic Progress." *Review of Austrian Economics*, 20(4): 221–36.

Walker, John. 2013. "U.S. Tax Code On-Line." Fourmilab. Accessed November 13.

Yamakawa, Yasuhiro, Mike Peng, and David Deeds. 2010. "Revitalizing and Learning from Failure for Future Entrepreneurial Growth." *Frontiers of Entrepreneurship Research*, 30(6): 1–11.