

Serial Entrepreneurship: Learning by Doing?

Francine Lafontaine*

Kathryn Shaw**

March 20, 2014

*William Davidson Professor of Business Economics, Ross School of Business, University of Michigan

**Earnest C. Arbuckle Professor of Economics, Graduate School of Business, Stanford University

We thank Robert Picard for his invaluable data management and programming assistance and Jennifer Cryer for her excellent research assistance. We thank seminar participants at Columbia University, University of Stockholm, and Stanford University for their comments, as well as Charles Brown, Casey Ichniowski, Edward Lazear, and James Spletzer for helpful discussions.

Abstract

Among typical entrepreneurs, is the serial entrepreneur more likely to succeed? If so, why? We answer these two questions using a comprehensive and unique data set on all establishments started at any time between 1990 and 2011 to sell taxable goods and services in the state of Texas. An entrepreneur is defined as the owner of a new business. A serial entrepreneur is one who opens repeat businesses. The success of the business is measured by the duration over which the business is in operation. The data show that serial entrepreneurship is relatively uncommon in retail trade. Of the almost 2.3 million retail businesses of small owners of new businesses in our data, only 25 percent are started by owners who have started at least one business before, and only 8 percent are started by an owner who is still operating at least one other business started earlier. However, once one becomes an entrepreneur for a second time, the probability of becoming one a third time, or fourth time, and so on, keeps rising. Moreover, we find that an owner's prior experience at starting a business increases the longevity of the next business opened, and that controlling for person fixed effects, prior experience still matters. Finally, experience at starting retail businesses in other sectors (e.g. a clothing store versus a repair shop) is beneficial as well, though not as much as same sector experience, and not in the restaurant sector. We conclude that prior experience imparts general skills that are useful in running the new business.

Much has been said about the serial entrepreneur, or the entrepreneur who starts one business after another. The popular press has saluted the serial entrepreneur, suggesting that success rates are higher for serial entrepreneurs than for first time entrepreneurs (e.g. Inc. Magazine’s “11 Historic Serial Entrepreneurs” at <http://www.inc.com/ss/12-historic-serial-entrepreneurs>). Using data from venture-backed firms, Gompers, Kovner, Lerner and Scharfstein (2010) show that entrepreneurs with a track record of success are more likely to succeed in further endeavors than first-time entrepreneurs. The circumstances surrounding the success of the set of entrepreneurs that are the subject of their study, however, may have much to do with the funding they receive, rather than the talent or experience of the entrepreneur.¹ Among typical entrepreneurs, is the serial entrepreneur more likely to succeed? If so, why?

We seek answers to these two questions using a comprehensive and unique data set on all businesses that collect sales taxes from their customers, i.e. all establishments that sell taxable goods and services in the state of Texas. The data is a longitudinal dataset that follows this population of businesses as they are opened and closed in Texas from 1990 to 2011. An entrepreneur is defined as the person or entity that owns a new business. A serial entrepreneur is one who opens repeat businesses. The success of the business is measured by the duration over which the business is in operation. To focus on the success of entrepreneurs rather than that of large established firms, we limit our sample to owners who open new businesses between 1990 and 2011. We restrict the data further to only those owners who do not open more than twenty establishments.

A look into the extent and success of serial entrepreneurship in retail is appealing because retail entrepreneurship is so prevalent. When an entrepreneur is measured as one who runs or starts a business, retail firms are second in the list of industries for entrepreneurs. Using CPS data, the industry rankings for the self-employed are construction, retail trade, professional services, business services, and real estate. Technology-based firms are in the minority. This CPS measure of entrepreneurship would understate the technology driven firms, but other measures show a similar pattern. In all cases, technology firms are in the minority, and venture backed technology firms are a trivial percent of entrepreneurs.² The entrepreneur with the successful venture-capital backed firm may produce far more jobs than the typical entrepreneur, but there are relatively few of these venture-

¹ As identified in Gompers, Kovner, Lerner and Scharfstein (2010), there is evidence from their study and from Sorensen (2007), Kaplan and Schoar (2005), and Lu (2007) that companies that are funded by top venture firms are more likely to succeed.

² See Lazear (2005) page 662 for this point.

capital backed entrepreneurs. Understanding the extent of serial entrepreneurship outside of venture-capital based industries, in those areas where most entrepreneurship occurs, is important. Moreover, as we show below, the retail and small-scale service businesses that we focus on have very high churning rates. Providing insights into factors that can enhance their survival is valuable, for the entrepreneurs themselves and their employees.

The data show that serial entrepreneurship is relatively uncommon in retail trade. Of the almost two million owners that we follow over the two decades of our data, about 25 percent start more than one establishment. However, once one becomes an entrepreneur for a second time, the probability of becoming one a third time, or fourth time, and so on, keeps rising.

We find that serial entrepreneurs are considerably more successful. Success is measured as the number of days that the firm stays in business. The average duration of businesses in our data is about 1218 days, or about 40 months.³ The median duration is much shorter, however, at 730 days, or 24 months. A Weibull proportional hazard rate model shows that the probability of exiting from business falls with past experience at starting businesses: for owners with one or more past businesses, the probability of exit is seven percent lower than for those with no prior business opened. Those with past businesses exit at lower rates, and thus have longer durations in business.

The time period that is studied – from 1990 to 2011 – permits us to examine whether recessions change patterns of entrepreneurship, particularly for serial entrepreneurs. They do. Those establishments founded during recessions are more likely to survive. Apparently the trial of founding in tough times increases success rather than decreasing it. Alternatively, owners only start particularly promising businesses at those times, postponing more risky ventures to later. However, there is less gain for the serial entrepreneur. The serial entrepreneur, moreover, has somewhat more of an advantage in future success when he founds his establishment in tough times.

The second question above was, why does serial entrepreneurship raise success? The answer is that it imparts skills. After controlling for person fixed effects, experience matters.

The paper is organized as follows. Section I describes the framework that we rely on, which focuses on what skills lead to entrepreneurship. The next section describes the data set. That is followed in

³ Average and median durations reported here, and in Table 3, are underestimated because there is right censoring in the data, i.e. we do not observe full spells for those businesses still opened at the end of 2011.

Section III by the evidence of patterns of entrepreneurship in the Texas retail data. Section IV focuses on estimating the impact of past entrepreneurship on new business ventures, and Section V asks why serial entrepreneurs are more successful. Section VI concludes.

I. Theoretical Framework: Why is Serial Entrepreneurship Successful?

Why might serial entrepreneurs be more successful? The answer to this question is motivated by Lazear's (2005) work on what makes an entrepreneur.

Lazear (2005) theorizes that an entrepreneur is a "jack-of-all-trades," or a "generalist." A generalist is one who need not excel in any one skill, but who is competent in many. Why is a generalist of particular value in entrepreneurship? The entrepreneur must marshal resources from many different areas, and he must be able to identify the creative talents of employees in many different arenas. In contrast, the specialist is one who can work for others, others who have the talent to spot his skills and to combine people of different skills. The generalist entrepreneur gets the value that corresponds to the minimum of his return across his different skills; the specialist employee gets the maximum value from his greatest skill.

When Lazear takes his model to data, he finds clear support for his model. The data that he uses is from an alumni survey from the Stanford Graduate School of Business. This dataset has two remarkable features. The first feature is that he can define an entrepreneur as one who was "Founder – among those who initially started the business." Though the sample is clearly limited by its focus on MBAs only, it has the advantage of allowing for a definition of entrepreneurship that is more appealing than most. The second feature of his dataset is that he knows the background characteristics of the entrepreneurs.

Lazear hypothesizes and finds that MBA graduates who have a variety of roles in firms are more likely to become entrepreneurs. In addition, those who took a greater variety of courses as MBA students are more likely to become entrepreneurs.

A final key test addresses why generalists are more likely to become entrepreneurs. There are two interpretations. One is that generalists are simply endowed with the skill set that makes them better entrepreneurs and their varied labor market roles reflects this endowment. A second, alternative, explanation is that the experience in many different roles results in the skills that make the individual a

better entrepreneur. The empirical test to distinguish these two interpretations uses information about the individual's future roles: if future roles – after the entrepreneurship spell – increase the probability of entrepreneurship, then endowment is what matters. The data support the view that experience in different roles enhances entrepreneurship: future roles have no impact on the probability of entrepreneurship, but past roles remain highly significant.

These results suggest two hypotheses for why serial entrepreneurs are more successful, and thus why we should find this to be true also in our retail trade data. The first is analogous to the test for general skills as gained through endowment or through experience. Serial entrepreneurs may be more successful because they are endowed with the skills of an entrepreneur and this is reflected in their higher success rates. Or, serial entrepreneurs may be more successful because the experience of entrepreneurship increases their knowledge base.

To our knowledge, most prior research on serial entrepreneurship has been on tech entrepreneurs.⁴ Because it is about tech firms, much of the work is on the interaction between serial entrepreneurship and funding options. Venture capital firms play a very limited role in the types of industries on which we focus. Instead, owners finance their businesses with more traditional sources of funds, including family and bank loans. Thus, our study complements the existing literature whose focus has been on VC-financed firms, and more specifically firms in high-tech industries.

II. Data

The data is the universe of all establishments that must collect sales taxes from their customers, i.e. all retail and personal service businesses, founded in the state of Texas from 1990 through 2011. These data were obtained by downloading public datasets of all sales tax-paying establishments in the state. We limit our data to businesses operating in the retail sector, broadly defined. Specifically, given that the businesses in our data are known to collect sales taxes, we include those that report doing business in the wholesale or retail sector, as well as those in the accommodation and food services sector, the Real Estate and Rental and Leasing sector, and some others.⁵ After grouping records for the same business, eliminating some sales-tax collecting entities that are not businesses, and reducing to only those businesses started after January 1, 1990, we have data for 2,780,370 establishments founded

⁴ On serial tech entrepreneurs, see Gompers, Kovner, Lerner and Sharfstein (2006), Ennew, Robbie, and Wright (1997), Hsu (2007), and Zhang (2011). For models of repeat entrepreneurship, see Amaral, Baptista, and Lima (2011), Kuechle, Menon, and Sarasvathy (2013), and Ucbasaran, Westhead, and Wright (2009).

⁵ Table A1, in the appendix, provides a full list of 3-digit NAICS that we include in our definition of retail.

between 1990 and 2011, by 1,715,352 separate owners.⁶ The Data Appendix provides further details on the data and sources.

For each establishment, we know the opening and closing dates, the name of the business, the name of the owner, the type of owner (proprietor, partner, or corporation), the industry code (SIC or NAICS, depending on the year), and whether the establishment is in an urban area (within city limits).⁷

For our regression analyses below, we exclude from the data those owners who opened more than twenty establishments. The reason for this limit is that we wish to study the opening and survival of stores or restaurants by small entrepreneurs, not the opening of large chain stores. This restriction eliminates only 2240 owners with a total of 120,323 establishments, a small loss compared to the remaining population of establishments. Our final sample for regression analyses relates to 2,331,998 businesses and 1,713,112 owners.

Some of the small business owners that remain in our data nonetheless open businesses under a national brand, as franchisees. These are included in our analyses. We have matched the names of the businesses to assess whether they are a part of such a chain, and less than 2 percent of the businesses of small business owners, i.e. those with 20 or fewer establishments, are associated with a national brand.⁸

III. Patterns of Retail Entrepreneurship

Entrepreneurs are defined as those owners that open establishments, whether they are organized as corporations, proprietorships, or partnerships. Table 1 shows that the number of stores and service establishments in existence in a given year, which we calculate by identifying all those in existence on July 4th of that year, whether they were started before or after our data period. As shown in Table 1, the population of retail and service establishments opened in Texas from 1990 to 2011 is very large,

⁶ Unfortunately, we do not observe the level of taxes paid, and thus cannot infer revenues for these businesses.

⁷ The births and deaths of establishments are defined by the time at which they report their tax obligations. Alternative databases on births and deaths would produce somewhat different results, according to the work of the BLS and of Spletzer (2000).

⁸ We used several directories to identify large national chains in the restaurant, retail, and personal and repair services, and in franchising. We searched for establishments of these chains using name matching. We also identified some chains directly by searching for business names that occurred often in the data. In total, we looked for outlets operating under about 1000 different brand names, and found that about 700 of these chains had operations in Texas. Because, as described further below, we eliminate all owners with more than 20 establishments, we are excluding a number of large franchisees as well as fully corporate chains from our data. This is why our proportion of chain stores is so low, much below the 10 or so percent of employer establishments that have been identified as belonging to franchised chains in the 2007 Census. Our results, however, are not sensitive to the exclusion of large owners from our data.

with 544,377 retail establishments still in existence in 2011.⁹ The population of establishments grew by more than 50% over the time period of our data. This is not surprising given that the population of people in Texas also grew by that much between 1990 and 2011, from 17,044,714 to 25,674,681 according to the Census.

Table 1: Number of Retail Establishments, and Entry and Exit, by Year

Year	Total establishments on July 4th	Establishments Opened Jan. 1 to Dec. 31	Establishments Closed Jan. 1 to Dec. 31
1990	362218	96798	65825
1991	398044	113171	75466
1992	448176	138925	100373
1993	476462	137260	126442
1994	480257	131263	107545
1995	495524	129451	132853
1996	493850	115859	111184
1997	495244	108357	98381
1998	508433	111272	123683
1999	501515	113882	119866
2000	492420	105998	101733
2001	493239	106183	96218
2002	511594	119080	107765
2003	524427	119918	103943
2004	535528	112690	119264
2005	527771	106480	123905
2006	511260	106678	103753
2007	510715	105125	101933
2008	509918	91976	87467
2009	514593	92125	91503
2010	519145	98492	79264
2011	544377	91328	36083
Total		2,452,311	2,214,460

Column 1 shows the total number of establishments as of July 4 each year. Columns 2 and 3 show the entry and exit of retail establishments, where retail is defined broadly, as detailed in the Data Appendix, each year.

⁹ According to the US Census' County Business Pattern data, in 2011, in the state of Texas, there were 326,105 establishments in the set of 2-digit NAICS that are part of the broad retail sector defined here (author's calculation of the sum of number of establishments in Texas at <http://censtats.census.gov/cgi-bin/cbpnaic/cbpsect.pl>). This number of establishments includes those of 3-digit NAICS that we do not include here, yet our number of establishments is greater. This is because the County Business Pattern data is restricted to employer establishments whereas our data include both non-employer and employer establishments. Unfortunately, we do not have a measure of employment, so we cannot separately identify employer and non-employer businesses.

The level of entry and exit each year is most striking in Table 1. The data on the rates of new firm creation and exit is incomplete in 1990 and 2011 (our 2011 data, for example, really ends with firms started in November). But in total, during the time period of our data, 2,452,311 establishments were opened, and 2,214,460 were closed. Clearly the relatively monotonic growth in the total number of establishments in Column 1 hides a substantial amount of churning in these industries – with about 20 percent of establishments exiting each year – a level of churning that corroborates the relatively short duration of businesses described above and further below.

The vast majority of new retail establishments are not opened by large chains. In Table 1 above, there are 2,452,311 new establishments opened in the state from 1990 through 2011. If we restrict our data to those owners who open 20 or fewer establishments, the number of newly opened establishments falls only by 120,323. In other words, despite the well-documented growth of chains in retail trade (notably Jarmin, Klimek, and Miranda, 2012, Basker, Klimek and Hoang Van, 2012, Cardiff-Hicks, Lafontaine and Shaw, forthcoming) large owners open less than 5 percent of new establishments in our data.¹⁰ Table 2 describes the number, entry, and exit for the subset of establishments that belong to large owners in our data.

To study entrepreneurial firms in the majority of this paper, we restrict the population of businesses to those of owners who open twenty or fewer establishments.¹¹ The aim of this restriction is to study the opening and closing of new establishments by entrepreneurs, not by large chains. However, in the last section of the paper (section VI), we estimate models for the large firms that operate the set of establishments described in Table 2, to contrast the results with those results from entrepreneurial retail owners.

¹⁰ There are many reasons for the smaller number of large owners in our data. First, our definition of retail is broader than that used in most papers, and this affects how comparable the data are. Second, as mentioned earlier, our data include a large number of non-employer businesses. Third, while franchisees would be counted as part of large chains in most papers, in our data it is the owner that matters, and so franchisees that own fewer than 20 establishments are among our small owners, no matter the size of the chain they are affiliated with.

¹¹ As described in the data appendix, the 20 establishments' cutoff is assessed including those that were opened prior to 1990. That way, a large chain that had many establishments at the beginning of our data period but few establishments opened in the 1990-2011 time frame will still be excluded from our study of new entrepreneurs.

Table 2: Number of Retail Establishments of Large Owners, and Entry and Exit, by Year

Year	Total establishments on July 4th	Establishments Opened Jan. 1 to Dec. 31	Establishments Closed Jan. 1 to Dec. 31
1990	32856	3875	3421
1991	33116	4253	4526
1992	33725	5085	3871
1993	33946	4700	4357
1994	34706	4624	3214
1995	36417	5841	3838
1996	37584	4722	4305
1997	38995	5784	4555
1998	39638	6410	5805
1999	41494	7505	5593
2000	42880	6229	4563
2001	44643	7435	6182
2002	45149	5531	5236
2003	45463	5169	4039
2004	46896	5066	4125
2005	48046	5640	3753
2006	49970	5879	4257
2007	51354	8480	7348
2008	52566	5041	3331
2009	53759	4680	3788
2010	54600	4305	3539
2011	55479	4069	2346
Total		120,323	95,992

Column 1 shows the total number of establishments as of July 4 each year. Columns 2 and 3 show the entry and exit of retail establishments, where retail is defined broadly, as detailed in the Data Appendix, each year.

What is striking in the data is that retail entrepreneurship, at least for small business owners, tends to be a single establishment affair. Table 3 shows the number of businesses that are started by owners who have not yet started any business before, and by those who started one before, and those who had started two and so on. It illustrates how the vast majority of small businesses are started by owners with no or very limited prior experience. Within the set of establishments owned by owners with no more than 20 in total, only 25.6 percent (namely 100 – 74.4 percent) of the businesses are owned by individuals who had opened another establishment since 1990 by the time they opened the focal one. And only 9 percent of them belong to owners who had opened two or more establishments by the time they opened a new one. So in general one should keep in mind that when we model serial

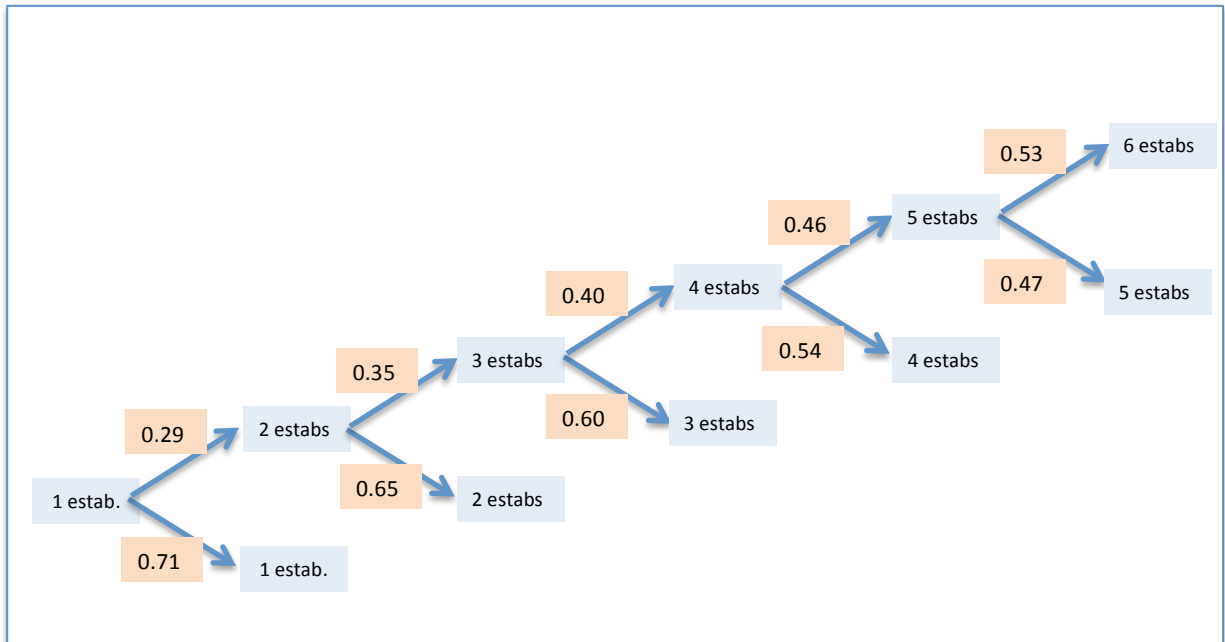
entrepreneurship empirically in this sector, we mostly model the impact of owning only one or two previous establishments. In fact, because most owners do not own multiple establishments, the number of owners in Texas is huge. As mentioned above, the 2.3 million new establishments in our data are run by 1,713,112 separate owners. In other words, the retail and small-scale service sector that our data cover affects the livelihood of a very large number of small business owners. Finally, the last two columns in Table 3 show that most of the owners of these establishments do not operate them concurrently but operate them sequentially. In this case, only eight percent of establishments are operated by an owner who has one or more establishments currently in operation. In this sector, entrepreneurship is mostly sequential.

Table 3: Number of establishments founded after Jan. 1, 1990, by number founded before by the owner, and by number still open, for small owners (≤ 20 outlets)

Owner's Prior Number of Establishments	Number of Establishments of Owners with Such Number of Prior Establishments	%	Number of Establishments of Owners with Such Number Founded Before that are Still Open	%
0	1,734,407	74.37	2,146,616	92.05
1	383,837	16.46	131,088	5.62
2	118,382	5.08	24,994	1.07
3	44,567	1.91	10,519	0.45
4	19,900	0.85	5,890	0.25
5	10,342	0.44	3,726	0.16
6	6,094	0.26	2,611	0.11
7	4,057	0.17	1,889	0.08
8	2,676	0.11	1,313	0.06
9	2,016	0.09	1,000	0.04
10	1,510	0.06	708	0.03
11	1,174	0.05	504	0.02
12	865	0.04	337	0.01
13	685	0.03	281	0.01
14	511	0.02	209	0.01
15	381	0.02	153	0.01
16	271	0.01	76	0
17	179	0.01	50	0
18	97	0	17	0
19	37	0	7	0
Total	2,331,988	100	2,331,988	100

Another interesting feature of serial entrepreneurship is that the probability of opening an additional establishment rises with the number of prior businesses opened. This is shown in Figure 1, where we examine the probability that an owner who has opened some number of businesses will open yet more of them, all this calculated within 15 years from their first business opened. Figure 1 shows that given one establishment opened in our data period, i.e. since 1990, the probability of opening a second one or more within the next 15 years of the first is 29 percent. But given two establishments opened, the probability of opening a third is 35 percent, and given three, the probability of opening a fourth is 40 percent. In other words, the probability of further repeat entrepreneurship rises with the number of prior businesses opened.

Figure 1: Extent of Serial Entrepreneurship



Note: Calculated for small owners in retail (those that will have no more than 20 establishments), for consistency with other analyses, and restricting to establishments opened after 1990, and within fifteen years from the owner's first observed establishment (to control for time in sample).

Table 4 shows descriptive statistics for the population of 2,331,988 retail businesses of small owners (defined as less than 20 establishments in our initial data) opened from 1990 to 2011.

The businesses in our data have an average duration of 1218 days, or about 42 months. The median duration is much smaller, however, at 730 days, or 24 months.¹² These average and median duration underestimate the true duration of the businesses, because the data are right censored, namely there are

¹² As described in more details in the data appendix, we limit the data to those businesses that survive at least 30 days. This explains the minimum duration shown in Table 4.

some businesses in our data that are still opened at the end of our data period and for which we do not observe a complete spell. In calculating the mean and median reported here, the duration of these establishments is counted as if they closed on the last day of our data, at the end of 2011.¹³

The duration models below will model exit as a function of the prior experience of the current owner. Prior experience is measured as the number of “businesses opened before,” as described in Table 4, which can either be “still open” or “now closed” (see rows 3 and 4). Table 4 shows that on average, the owners of the businesses have opened only 0.45 establishments by the time they start a new business. Consistent with information shown in Table 3, the majority of the prior businesses, or 68.8 percent of them, are closed by the time the new business is opened ($.313/.455 = 68.8\%$). Finally, owners start as many businesses after the focal one as they do before ($.460/.455$).

Table 4: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Duration (in days)	1217.721	1359.501	30	8034
Businesses opened before: all	.455	1.133	0	19
Businesses opened before: still open	.142	.706	0	19
Businesses opened before: now closed	.313	.776	0	19
Businesses opened after current business	.460	1.105	0	19
Urban establishment	.822	.382	0	1
National chain	.020	.139	0	1
Opened in recession	.104	.305	0	1
Corporation	.220	.414	0	1
Proprietorship	.687	.464	0	1
Partnership	.093	.290	0	1
Survives 1 year (n = 2244729)	.723	.448	0	1
Survives 2 years (n = 2150542)	.533	.499	0	1
Survives 3 years (n = 2063097)	.411	.492	0	1

Number of observations: 2331988 except as noted. The number of observations used to calculate the survival rates is reduced to the set of businesses that start one, two or three years prior to the end of our data period to ensure we can observe a full year, two-year or three-year survival.

Table 4 also shows that the vast majority of the businesses are in urban settings, and only a small proportion of these are associated with national chains, as mentioned earlier. The latter is to be expected as our goal of examining the extent of entrepreneurship among small business owners led us to eliminate from the data the businesses of larger owners. This means that establishments of large

¹³ An alternative way to handle this issue would be to exclude all establishments that are still opened at the end of our data period. Accounting for right censoring is a main reason to use duration models in our analyses below.

franchisees and those of all large corporate chains with significant presence in Texas are excluded from the data.

Interestingly, 10.4 percent of establishments are founded during recession months. This is somewhat below the 11.7 percent of all months covered by our data that are recession months, which suggests that the rate of retail business creation is slower during such months.¹⁴

In this paper, we label the owners of businesses established during our data period as entrepreneurs. Most of these are proprietors: 69 percent of the businesses in our data are owned by proprietors. Second to that are corporations: these own 22 percent of the businesses in our data, leaving less than 10 percent owned by partnerships.

IV. Is Serial Entrepreneurship More Successful?

The primary hypothesis we want to investigate is that serial entrepreneurs will have higher success rates than first-time entrepreneurs. Models that emphasize the importance of learning-by-doing for example would suggest that the businesses of those that have started businesses before will be more successful. In the subsections below, we examine how much serial entrepreneurship matters for business success, and why it matters.

A. The Value of Serial Entrepreneurship

Entrepreneurship is said to be more successful when the establishment stays in business longer. Therefore, the goal is to measure the impact of the number of past businesses opened on the duration of a focal business, where duration again is measured in days since opening.

In this section, we explore these factors using duration models. Two models are estimated: the Cox model and the Weibull model. The Cox proportional hazards model has the advantage of not relying on any distributional assumption, and as such provides a useful robustness check. The main drawback of the Cox model for our purposes is that it assumes a constant hazard whereas the literature on firm survival suggests that negative duration dependence is to be expected in the data. We therefore also show results from a Weibull model, which can capture negative duration dependence in the data. Also, if the Weibull distribution fits the data well, the estimates from the Weibull are more efficient.

¹⁴ Fairlie (2011) finds that more entrepreneurship occurs during periods of high unemployment (recession). The difference may be attributed to the construction trade, which accounts for an important portion of the newly self-employed in his data, and is a sector within which he finds that a sizable number of people turn to self-employment during downturns.

In the Weibull model, the hazard, or instantaneous transition from origin (start of business) to destination (business exit) given that the business has survived to time t , can be written as $h(t) = h_0(t)g(X)$, where $h_0(t) = pt^{p-1}$ and p is the shape parameter. In this hazard function, if $p < 1$, we have negative duration dependence, meaning that older businesses have lower exit rates, whereas $p > 1$ indicates positive duration dependence. When $p = 1$, the Weibull model reduces to the Cox model.

The Cox and Weibull models both exhibit the “proportional hazard rate” property, i.e. their hazard function, which is the rate at which a chain exits given it has survived until time t , can be written as $h(t) = h_0(t)e^{X'\beta}$. Changes in regressors thus shift the baseline hazard, $h_0(t)$, and the exponentiated coefficients capture the effect of a one-unit increase in a particular variable on the hazard ratio. Specifically, if the exponentiated coefficient b is greater (smaller) than one, the difference $(b-1)*100$ indicates the percentage by which a one unit increase in the explanatory variable would increase (decrease) the hazard of exit.¹⁵ For that reason, we report exponentiated coefficients in the tables below, so that a reported coefficient that is greater than one indicates that the variable increases the exit hazard rate, while a variable with a coefficient below one reduces it. The reported standard errors are clustered at the owner level. Also, as is standard in this type of estimation, the levels of significance – as indicated by stars in the table – are assessed based on original coefficients and standard errors.

Two further issues affect our estimations. First, our data end in 2011. For businesses that exit before that date, we observe their end-of-business date and thus we know their full duration spell. For non-exiting businesses, however, the duration spells are incomplete and their observations are right censored. The second issue is that a business started prior to 1990 could only be present in our sample if it survived at least until 1990. If a business started prior to 1990 did not survive to that time, we would not know it ever existed. Thus, for owners of businesses started prior to 1990, there would be a survivorship bias in our counts. For that reason, as described earlier, we focus on those businesses that owners started since 1990 in all our analyses below. None of our results, however, are sensitive to this restriction.

In the data, each retail outlet can be operated by an owner who has had up to nineteen previous stores. Therefore, the first specification models the exit as a function of nineteen dummy variables for the number of previous stores, where these dummies are “1 business before” to “19 businesses before” (the omitted category is “0 business before”).

¹⁵ Suppose that we have only one covariate, X , that we increase by 1 unit. The ratio of exit hazards after and before this change can be expressed as a function of the coefficient of X , namely: $\frac{h(t|X+1)}{h(t|X)} = \frac{h_0(t)e^{(X+1)\beta}}{h_0(t)e^{X\beta}} = e^\beta$.

Table 5: Full Sample Duration Regressions: Dummy Variable Specification

Owner Opened:	Cox	Weibull
1 business before	0.917*** (0.002)	0.928*** (0.002)
2 businesses before	0.894*** (0.003)	0.912*** (0.003)
3 businesses before	0.899*** (0.005)	0.923*** (0.006)
4 businesses before	0.922*** (0.008)	0.950*** (0.009)
5 businesses before	0.933*** (0.012)	0.963*** (0.013)
6 businesses before	0.925*** (0.016)	0.953*** (0.017)
7 businesses before	0.939*** (0.021)	0.971 (0.023)
8 businesses before	0.917*** (0.025)	0.953* (0.028)
9 businesses before	0.939* (0.030)	0.975 (0.033)
10 businesses before	0.872*** (0.035)	0.908** (0.039)
11 businesses before	0.915* (0.042)	0.957 (0.046)
12 businesses before	0.937 (0.045)	0.983 (0.049)
13 businesses before	0.978 (0.054)	1.022 (0.059)
14 businesses before	1.018 (0.072)	1.066 (0.078)
15 businesses before	0.965 (0.074)	1.014 (0.081)
16 businesses before	1.064 (0.102)	1.132 (0.111)
17 businesses before	1.113 (0.119)	1.188 (0.129)
18 businesses before	1.389** (0.186)	1.500*** (0.204)
19 businesses before	1.945*** (0.372)	2.125*** (0.394)
Opened in recession	0.893*** (0.002)	0.918*** (0.003)
National Chain	0.744*** (0.006)	0.730*** (0.007)
Urban establishment	1.208*** (0.002)	1.212*** (0.003)
Proprietorship	1.644*** (0.004)	1.677*** (0.004)
Partnership	1.710*** (0.006)	1.752*** (0.007)
Number of obs.	2,331,988	2,331,988
No. of failures	1,849,592	1,849,592
p (Weibull)	.	0.88

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Results in Table 5 indicate that past entrepreneurship has a strong negative effect on the probability of closing of the current business. In column 1 are the coefficients for the Cox model. Looking across the coefficients on the “1 business before” to “19 businesses before” dummy variables, it is clear that past entrepreneurship lowers the exit rate, except when the number of businesses before becomes very large. We come back to this below. In column 2 are the Weibull results. The Weibull is the better model; the probability of exit falls with duration in the state according to the estimated shape parameter, p , at the bottom of the table, which we find to be statistically significantly less than 1. Therefore, in subsequent tables, we report only Weibull results. The pattern of coefficients of the “1 business before” to “19 businesses before” variables are the same as for the Cox model. For those owners with one previous business, the probability of exit for the current business falls by 7 percent ($1 - .928$).

The dummy variables for the number of businesses opened before the focal business suffer from a flaw – as the number of prior businesses rises, the dummy variables become little populated. As shown in Table 3, only 25.6 percent of new businesses are started by owners that have previous experience with starting businesses. Only 4 percent of businesses are opened by owners who have opened three or more before the current one. Therefore, the dummy variables in Table 5 are little populated after three past businesses. A better way of capturing experience effects may be to aggregate the 19 dummy variables into a linear variable for number of past businesses.

We show results from introducing the number of previous businesses opened linearly in columns 1 and 5, and in a quadratic way in columns 2 and 6, of Table 6. Results imply that the number of businesses opened prior to the focal one has a positive effect on its duration, though at a diminishing rate (because the squared term is positive in columns 2 and 6).¹⁶ More experience in serial entrepreneurship matters: Current business duration increases as the number of prior businesses opened rises. This effect of past businesses on current success was not evident in the coefficients on the nineteen dummy variables; these exponentiated coefficients did not fall as would be expected. But again, the dummy variables representing more than three past stores are not very populated.

Recall that we can ascertain from the data whether businesses opened in the past remain in business or are now closed. The literature suggests there are two paths to what we have referred to as serial entrepreneurship, which is also known as “habitual entrepreneurship” in some of the literature.¹⁷ In

¹⁶ In these models, and all those shown below, the results do not change when we introduce dummy variables for NAICS industry codes. Results from estimating the same regressions as in Table 6 with 3-digit NAICS industry fixed effects are shown in Appendix B.

¹⁷ For these definitions, see Westhead and Wright, 1998 and Birley, 1993. Consistent with Lazear (2005), and much of the trade literature, we use serial entrepreneur to mean someone who has opened more than one

one case, the entrepreneur opens and then closes a series of businesses in sequence, typically operating only one at any given time. In a second, labeled portfolio entrepreneurship in some of the literature, the entrepreneur opens and keeps running a number of businesses at the same time.

Table 6: Weibull Duration Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Number Opened	0.980***	0.954***			0.957***	0.916***
Before	(0.001)	(0.002)			(0.003)	(0.003)
Squared (Number		1.004***				1.009***
Opened Before)		(0.000)				(0.001)
Number Opened			1.003	1.017***		
Before Still Open			(0.003)	(0.004)		
Sq. (Number Opened				0.998***		
Before Still Open)				(0.001)		
Number Opened			0.962***	0.924***		
Before but Closed			(0.001)	(0.002)		
Sq. (Number Opened				1.011***		
Before but Closed)				(0.000)		
Opened in recession	0.919***	0.919***	0.919***	0.919***	0.922***	0.920***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Major Chain	0.735***	0.733***	0.726***	0.724***	0.724***	0.724***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Urban Establishment	1.211***	1.212***	1.210***	1.210***	1.200***	1.198***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Proprietorship	1.672***	1.675***	1.682***	1.685***	1.672***	1.675***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Partnership	1.755***	1.752***	1.757***	1.757***	1.755***	1.753***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Number Opened					0.992***	1.002
Before * Recession					(0.003)	(0.005)
Number Opened					1.018***	1.028***
Before * Chain					(0.006)	(0.010)
Number Opened					1.027***	1.046***
Before * Urban					(0.003)	(0.004)
Sq. (Number Opened						0.998***
Before * Recession)						(0.001)
Sq. (Number Opened						0.997**
Before * Chain)						(0.001)
Sq. (Number Opened						0.995***
Before * Urban)						(0.001)
Number of obs	2,331,988	2,331,988	2,331,988	2,331,988	2,331,988	2,331,988
No. of failures	1,849,592	1,849,592	1,849,592	1,849,592	1,849,592	1,849,592
p (Weibull)	0.88	0.88	0.88	0.88	0.88	0.88

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

We know from Table 3 that there is relatively little portfolio entrepreneurship in the retail sector. The vast majority of owners operate only one establishment at a time; entrepreneurship is sequential.

business, and we use sequential entrepreneurship to describe the version where businesses are opened and then closed such that the entrepreneur operates only one at a time.

But the natural question arises, do entrepreneurs with more than one store currently open fare better or worse than those with closed stores? To distinguish between open and closed stores, we replace the total number of businesses opened before by two variables, the “number opened before still open” for the number of stores that remain open, and the “number opened before but closed” for the number of past stores that are now closed. As shown in columns 3 and 4 of Table 6, the “number opened before but closed” is the significant variable in the regression, decreasing the exit rate of the current store. The “number opened before still open” variable has the reverse effect, but this effect is not statistically different from zero, and is much smaller in magnitude: in column 3, each additional business opened but now closed reduces the hazard rate by 3.8 percent (1 – 96.2). An additional business opened still operating increases the hazard by only .3 percent.

The conclusion is that for small business owners, serial entrepreneurship is most successful when it is truly sequential – when the stores are opened and closed and opened again sequentially rather than operated concurrently. This result could arise for multiple reasons. The closing of past stores may signify success if these past stores were sold to new owners. However, the conventional wisdom in the trade literature is that, because of information asymmetry problems, small business owners in the retail and restaurant and similar sectors often receive far less than the predicted future cash flows when they sell their businesses (e.g., Fraser, 1999). Retail trade is not like the tech business, wherein a new entrepreneur is especially successful when he sells off his business. Alternatively, the closing of past stores may teach the current owner how to be a better entrepreneur. Or, the closing of past stores may give the owner the time to focus on his current endeavor. Since most of the owners in the data are single proprietors, it is possible that running several businesses at once imposes too much of a burden on the small proprietor.

Note lastly that the form of ownership has a very significant effect on duration. The coefficients on partnership and proprietorship in our regressions above are substantially above one, confirming that indeed corporations are more successful on average at keeping businesses open.

The duration models estimated thus far make use of information on the number of businesses opened previously, but do not use information on the success of these businesses, other than whether they remain open. But we also know the duration of the past businesses. To investigate whether the duration of past businesses opened matters, we create a new measure of past experience, one that embodies the duration as well as the number of past businesses open. This is measured by the sum, across all previously started businesses, of the number of days that they each remain open, up to the date at which the new focal business is opened. We then divide this sum by the total number of days since the owner opened his first business. The measure thus represents the number of businesses that

the owner of the new business has operated, on average, since he/she started his/her first business. Not surprisingly, given that many businesses are closed before the new business is open, the average number of businesses that small business owners operate over the full period since starting their first business is lower than the number of new businesses started prior to starting the focal business. Specifically, the average number of businesses operated before is 0.274, to be compared to 0.455 in Table 4. The standard deviation of this alternative measure is also lower, at 0.642.

As shown in Table 7, when this variable replaces our usual measure of experience, we find that it also has a significantly positive effect on new business survival.

Table 7: Weibull Duration Regressions, Alternative Measure of Experience in Business

	(1)	(2)	(3)	(4)	(5)	(6)
Avg. Number	0.935***	0.902***	0.943***	0.913***	0.874***	0.841***
Operated Before	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.005)
Squared (Number		1.009***		1.008***		1.016***
Operated Before)		(0.001)		(0.001)		(0.001)
Number Opened			0.990***	0.974***		
Before but Closed			(0.002)	(0.003)		
Sq. (Number Opened				1.007***		
Before but Closed)				(0.000)		
Opened in recession	0.918***	0.918***	0.918***	0.917***	0.924***	0.923***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Major Chain	0.741***	0.739***	0.739***	0.737***	0.721***	0.721***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
Urban Establishment	1.213***	1.213***	1.213***	1.213***	1.194***	1.195***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Proprietorship	1.662***	1.663***	1.664***	1.666***	1.662***	1.664***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Partnership	1.748***	1.745***	1.748***	1.746***	1.748***	1.745***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Avg. Number Operated					0.974***	0.973***
Before * Recession					(0.006)	(0.007)
Avg. Number Operated					1.050***	1.072***
Before * Chain					(0.009)	(0.016)
Avg. Number Operated					1.077***	1.085***
Before * Urban					(0.006)	(0.007)
Sq. (Avg. Number Oper.						1.000
Before * Recession)						(0.001)
Sq. (Avg. Number Oper.						0.993***
Before * Chain)						(0.002)
Sq. (Avg. Number Oper.						0.993***
Before * Urban)						(0.001)
Number of obs	2,331,988	2,331,988	2,331,988	2,331,988	2,331,988	2,331,988
No. of failures	1,849,592	1,849,592	1,849,592	1,849,592	1,849,592	1,849,592
p (Weibull)	0.88	0.88	0.88	0.88	0.88	0.88

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

B. When is Serial Entrepreneurship Most Valuable?

In the hazard rate models of Tables 5 to 7, there are several control variables, and the one of greatest interest is “Opened in recession.” This is a dummy variable equal to one if the current business was opened during a recession, for the three recessions that occurred between 1990 and 2011.

Businesses opened during recessions are less likely to close. Apparently, these stores are more durable. Recall from Table 4 that only 10.4 percent of the stores are opened during a recession in these 22 years, within which 31 months (or 11.7 percent) are classified as recession months according to the NBER definition. It is therefore possible that the businesses opened during those periods are selected, i.e. that owners only pursue very promising business ideas during recessions.

Businesses opened by small business owners in urban areas, defined in our data source as businesses located within city limits, are more likely to be of shorter duration. In other words, it appears that more intense competition lowers success.¹⁸

Lastly, establishments of small business owners that are associated with a national brand are less likely to close. But keep in mind that the owner must have less than twenty businesses in total to be in our data set. This restriction explains why only 2.0 percent of the businesses in our data are part of chains. The businesses in question are owned by franchisees.

Each of these three variables is interacted with our measure of the extent of serial entrepreneurship in columns 5 and 6 of Tables 6 and 7. Serial entrepreneurship is more valuable to those stores opened in recessions, but less valuable to those stores in urban areas or those associated with national brands.

C. Is Entrepreneurship a General Skill?

In the model of Lazear (2005), entrepreneurs are those with general skills, not specific skills. The general skills he has in mind are from different management disciplines. An entrepreneur is said to be a generalist if he had multiple past roles from a history of working in operations, and finance, and human resources. An entrepreneur is said to be a specialist if he had few past roles, i.e. if he has a history of working in only one discipline, such as finance. Lazear’s data show that entrepreneurs tend to be generalists.

The next question that naturally arises is whether entrepreneurs tend to be generalists within the domain of entrepreneurship. Some famous technology entrepreneurs have opened firms in very

¹⁸ See also Asplund and Nocke, 2006, for a model and evidence that the lifespan of firms is shorter in larger markets.

different arenas.¹⁹ Perhaps more likely though is the entrepreneur who gets a series of ideas by working within one technology domain. For example, a group of techies from Google will spin off firms that are all technology based.²⁰

Our retail data enables us to test whether small business entrepreneurs should stick to one retail domain. Would a restaurateur do better today if he had gained experience running a clothing store in the past? Or should all his previous experience be in restaurants? This question is similar – though a bit different – to that asked by Lazear. If multiple past roles lead to more entrepreneurship, do multiple past business types or sectors of activity lead to better entrepreneurship? In other words, the entrepreneur’s past roles could include not just different managerial disciplines but also different product market arenas and production functions. For example, does the entrepreneur opening a restaurant benefit more if his past experience is in opening more restaurants, or does experience in all forms of retail and small-scale businesses raise success?

In Table 8, we return to our standard measure of prior business experience, namely the number of businesses previously opened, to explore whether past domain experience matters in the largest retail sub-industries in our data – those of restaurants (NAICS 722), clothing stores (NAICS 448), and repair businesses (NAICS 811). We also examine the general retail trade (NAICS 44-45) sector, where perhaps experience across types of retails may be useful, but experience outside retail may not be.²¹ Past experience is measured in two ways. The first is experience in the current domain, labeled “same sector,” measuring the number of past businesses within the same industry. The second form of experience is that in other domains, “other sectors,” measuring the number of past businesses outside the specific industry domain (defined at the 3-digit NAICS level, as described above, or at the relevant combination of 2-digit NAICS for the whole of retail trade). In each of these sectors, small business owners’ experience is concentrated in the same sector: in clothing, the proportion of same industry experience, measured by the number of prior businesses in this sector as a proportion of the total number of prior businesses opened by the owner, is .358/.455, or 78.7 percent on average. In the restaurant and repair sectors, the corresponding proportions are 78.8 percent (.394/.500) and or 73.7 percent (.264/.358) respectively.

Results in Table 8 show that in most sub-industries, experience in any industry raises success: both “same sector” businesses and “other sectors” businesses lower exit rates (columns 2 to 4). The

¹⁹ As an example, Elon Musk founded PayPal then SpaceX and he currently serves as CEO of Tesla Motors.

²⁰ This is known as “spawning” (Klepper, 2001, Klepper and Sleeper, 2005, Gompers, Kovner, Lerner and Sharfstein, 2005, and Chatterji, 2009).

²¹ Note that clothing stores are a subset of all retail trade businesses. We show results for this category separately as it is the largest of the subsectors in retail in terms of number of businesses in our data.

exception is restaurants, in column 1. In restaurants, it is really only past restaurant experience, i.e. experience in “same sector” that raises success. In restaurants, other forms of experience lower success. In the other sub-industries, any past experience raises success.

Table 8: Results by Sector

	Restaurants	Clothing	Repair	Retail
Bus. Opened Before	0.942***	0.941***	0.964***	0.957***
- Same Sector	(0.005)	(0.006)	(0.007)	(0.002)
Sq. (Bus. Opened Before	1.003***	1.006***	1.004	1.005***
- Same Sector)	(0.001)	(0.001)	(0.002)	(0.000)
Bus. Opened Before	1.042***	0.972***	0.972**	0.953***
- Other Sectors	(0.007)	(0.009)	(0.013)	(0.004)
Opened in Recession	0.942***	0.921***	0.914***	0.917***
	(0.008)	(0.009)	(0.010)	(0.004)
National Chain	0.578***	0.676***	0.880***	0.901***
	(0.009)	(0.057)	(0.037)	(0.014)
Urban Establishment	0.962***	1.257***	1.234***	1.231***
	(0.007)	(0.010)	(0.010)	(0.003)
Proprietorship	1.814***	1.573***	1.669***	1.640***
	(0.013)	(0.019)	(0.017)	(0.006)
Partnership	1.592***	1.796***	2.164***	1.770***
	(0.020)	(0.030)	(0.037)	(0.010)
Number of obs	248,938	181,546	146,982	1,286,416
No. of failures	191,466	151,519	111,530	1,060,777
p (Weibull)	0.83	0.90	0.83	0.90

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

D. The Duration of New Businesses: An Alternative Specification

An alternative approach to estimating the effect of serial entrepreneurship on the survival of businesses is to estimate a logit model for the probability of survival. The logit model is useful also because we can introduce fixed effects for owners, which we do in the next section. For now, we use this alternative modeling approach to verify the robustness of the above results.

We create a dummy variable equal to one if the focal business survives at least one year, and a different one for surviving at least two-year, and then for remaining in business at least three years from their opening date. Thus, there are three dependent dummy variables, SURVIVE1Y, SURVIVE2Y, SURVIVE3Y, for the three alternative survival durations. Note that we eliminate incomplete spells from these regressions, that is we limit the observations to those for which we can see that the business lasted at least 1, or 2, or 3 years respectively. This is the reason for the smaller sample sizes in these regressions.

The model to be estimated then is:

$$(1) \text{SURVIVEN}Y_{it} = b_0 + b_1 \text{NumberOpenedBefore} + b_2 \text{NumberOpenedBefore}_{it}^2 + \beta X_{it} + e_{it}$$

where NumberOpenedBefore again is measured by the number of businesses opened by the same owner before the current one. In these regressions, a positive coefficient indicates that the variable increases the likelihood that the business survives for the number of years specified. There are control variables X_{it} in the regression.

The results, in Table 9, further confirm those from our duration analyses above. They show that a greater number of businesses opened before the current one leads to a greater likelihood of survival across the three survival durations. Though the functional form is now the logit model instead of the Weibull model, prior experience as an entrepreneur influences current success.

Table 9: Logit Results: Dependent Variable is Survival for One, Two or Three Years

	1-year	2-year	3-year	1-year	2-year	3-year
Number Opened Before	0.030*** (0.002)	0.032*** (0.002)	0.034*** (0.002)	0.074*** (0.003)	0.082*** (0.003)	0.087*** (0.003)
Squared (Number Opened Before)				-0.007*** (0.000)	-0.008*** (0.000)	-0.008*** (0.000)
Opened in Recession	0.149*** (0.005)	0.174*** (0.005)	0.214*** (0.005)	0.149*** (0.005)	0.175*** (0.005)	0.216*** (0.005)
National Chain	0.539*** (0.018)	0.485*** (0.015)	0.478*** (0.015)	0.545*** (0.018)	0.491*** (0.015)	0.484*** (0.015)
Urban Establishment	-0.297*** (0.004)	-0.309*** (0.004)	-0.329*** (0.004)	-0.298*** (0.004)	-0.310*** (0.004)	-0.330*** (0.004)
Proprietorship	-0.781*** (0.005)	-0.757*** (0.004)	-0.769*** (0.004)	-0.784*** (0.005)	-0.760*** (0.004)	-0.772*** (0.004)
Partnership	-0.886*** (0.007)	-0.854*** (0.006)	-0.849*** (0.007)	-0.884*** (0.007)	-0.851*** (0.006)	-0.846*** (0.007)
Constant	1.815*** (0.006)	0.955*** (0.005)	0.476*** (0.005)	1.808*** (0.006)	0.946*** (0.005)	0.466*** (0.005)
N	2,244,729	2,150,542	2,063,097	2,244,729	2,150,542	2,063,097

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Sample sizes are reduced as we go from the one-year to two-year to three-year sample because we need to observe complete spells in these analyses. The sample is therefore limited to those businesses that are started prior to the end of 2010, end of 2009 and end of 2008 respectively in these survival analyses.

V. Why Is Serial Entrepreneurship More Successful?

Serial entrepreneurship may have value for two main reasons. It may be that the early experience of entrepreneurship is a period of learning and that this learning increases the future success of the entrepreneur's businesses. Alternatively, it may be that the entrepreneur with many businesses is just naturally or innately good at entrepreneurship and his first business is likely to be just as successful as any future business. In this case, success at serial entrepreneurship is just a proxy for high underlying entrepreneurial skills.

To identify which of these two scenarios is the source of success for serial entrepreneurs, we use the logit model again. The model to be estimated is:

$$(2) \text{SURVIVE}_{it} = b_0 + b_1 \text{NumberOpenedBefore}_{it} + b_2 \text{YearlyNumberOpenedAfter}_{it} + \beta X_{it} + e_{it}$$

where $\text{NumberOpenedBefore}$ is measured as always, as the number of businesses opened before the current one, from day one in our data, in 1990. The $\text{YearlyNumberOpenedAfter}$ is the number of stores opened per year after the current store. This variable is constructed by dividing the NumberOpenedAfter by the number of years left in the data from the start date of the focal unit, YearsLeft_{it} .

The maintained hypothesis is that serial entrepreneurs are more successful because entrepreneurship is a learned skill. In this case, future entrepreneurship will have no effect on current survival of a business, rejecting $b_2 > 0$, because entrepreneurial success depends only on the past experience of running businesses (implying $b_1 > 0$). The alternative hypothesis is that $b_2 > 0$ because future entrepreneurship is a proxy for unobserved talent. This model is analogous to that in Lazear (2005), where the probability of being an entrepreneur is a function of the number of job roles before and after the individual's entrepreneurship spell, after controlling for time in sample.

For all three alternative survival durations, of at least one to three years, we find that $\text{YearlyNumberOpenedAfter}$, i.e. the frequency with which the owner starts business after the focal one, has a negative effect on survival while $\text{NumberOpenedBefore}$ continues to have a positive effect (Table 10, columns 1 through 3). We further verify the robustness of this result using a specification where we control for YearsLeft and allow the NumberOpenedAfter to enter into the regression model in a quadratic way, as we do for $\text{NumberOpenedBefore}$ (columns 4 through 6). Again, we find a negative effect of future business openings, which seems to reinforce our earlier results about the sequential nature of serial entrepreneurship in these data. In other words, we again find evidence that owners close their current business before opening others, suggesting that one business crowds out the

other. However, prior experience continues to have a strong positive effect on the survival of the focal business. Prior experience appears to teach entrepreneurship.

Table 10: Logit Results: Dependent Variable is Survival for One, Two or Three Years

	1-year	2-years	3-years	1-year	2-years	3-years
Number Opened Before	0.092*** (0.003)	0.098*** (0.003)	0.102*** (0.003)	0.077*** (0.003)	0.090*** (0.003)	0.097*** (0.003)
Sq. (Number Opened Before)	-0.007*** (0.000)	-0.008*** (0.000)	-0.008*** (0.000)	-0.007*** (0.000)	-0.008*** (0.000)	-0.008*** (0.000)
Yearly Number Opened After	-0.651*** (0.023)	-0.599*** (0.024)	-0.591*** (0.027)			
Total Number Opened After				-0.105*** (0.003)	-0.118*** (0.003)	-0.131*** (0.003)
Sq. (Total Number Opened After)				0.012*** (0.001)	0.014*** (0.001)	0.015*** (0.001)
Years Left from Focal Bus. Start				-0.007*** (0.000)	-0.004*** (0.000)	-0.002*** (0.000)
Opened in Recession	0.156*** (0.005)	0.181*** (0.005)	0.221*** (0.005)	0.125*** (0.005)	0.157*** (0.005)	0.207*** (0.005)
National Chain	0.578*** (0.018)	0.518*** (0.016)	0.509*** (0.015)	0.553*** (0.018)	0.496*** (0.015)	0.489*** (0.015)
Urban Establishment	-0.293*** (0.004)	-0.305*** (0.004)	-0.326*** (0.004)	-0.281*** (0.004)	-0.298*** (0.004)	-0.319*** (0.004)
Proprietorship	-0.804*** (0.005)	-0.775*** (0.004)	-0.786*** (0.004)	-0.780*** (0.005)	-0.761*** (0.004)	-0.775*** (0.004)
Partnership	-0.906*** (0.007)	-0.869*** (0.007)	-0.862*** (0.007)	-0.892*** (0.007)	-0.865*** (0.006)	-0.864*** (0.007)
Constant	1.840*** (0.006)	0.973*** (0.005)	0.491*** (0.005)	1.918*** (0.007)	1.020*** (0.006)	0.531*** (0.006)
N	2,244,729	2,150,542	2,063,097	2,244,729	2,150,542	2,063,097

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Sample sizes are reduced as we go from the one-year to two-year to the 3-year sample as we need to observe complete spells in these analyses. The sample is therefore limited to those businesses that are started prior to the end of 2010, end of 2009 and end of 2008 respectively in these survival analyses.

An alternative means of estimating the maintained hypothesis is to use a logit regression that now includes owner fixed effects. In this case, the regression to be estimated is

$$(3) \text{SURVIVEN}Y_{it} = b_0 + b_1 \text{NumberOpenedBefore}_{it} + \beta X_{it} + \alpha_i + e_{it}$$

where α_i is an owner fixed effect and the functional form for (3) is the logit. If $b_1 > 0$ after controlling for innate talent α_i , then entrepreneurship is successful because it is in part a learned skill, not entirely an innate talent. Note that since the organizational form (proprietor, corporation, and partnership) is at

the owner level, the inclusion of owner fixed effects makes it impossible to continue to control for ownership form.

The results, in Table 11, support the maintained hypothesis: entrepreneurship is a learned skill. In this table, columns 1 to 3 present the standard logit results for the subsample of owners who own one or more businesses. We have reduced the data to this subsample so that we can subsequently run the fixed effects regressions.²² For this subsample of repeat entrepreneurs, NumberOpenedBefore has a much larger, and highly significant, effect. In the fixed effects logit regressions, in columns 4 through 6, the NumberOpenedBefore continues to have a large positive effect even within owners.

Table 11: Fixed Effects Logit Results
Dependent Variable is Survival for One, Two or Three Years

	1-year	2-years	3-years	1-year	2-years	3-years
Number Opened Before	0.296*** (0.004)	0.236*** (0.003)	0.224*** (0.004)	0.237*** (0.004)	0.288*** (0.004)	0.342*** (0.004)
Squared (Number Opened Before)	-0.020*** (0.000)	-0.019*** (0.000)	-0.020*** (0.000)	-0.023*** (0.000)	-0.028*** (0.000)	-0.033*** (0.001)
Opened in Recession	0.009 (0.010)	-0.009 (0.009)	0.077*** (0.010)	0.021* (0.011)	-0.037*** (0.010)	0.047*** (0.011)
National Chain	0.384*** (0.024)	0.329*** (0.020)	0.343*** (0.020)	0.434*** (0.045)	0.377*** (0.041)	0.349*** (0.042)
Urban Establishment	-0.171*** (0.009)	-0.238*** (0.008)	-0.282*** (0.008)	-0.253*** (0.011)	-0.329*** (0.011)	-0.383*** (0.011)
Proprietorship	-0.237*** (0.008)	-0.261*** (0.007)	-0.259*** (0.007)			
Partnership	-0.139*** (0.014)	-0.088*** (0.013)	-0.033** (0.013)			
N	432,981	486,975	445,867	432,981	486,975	445,867
Nb. Groups				151,264	171,718	158,013

Notes: Standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. In columns 1 to 3, standard errors are clustered at the owner level. In the fixed effects regressions, we cannot include proprietorship or partnership indicator variables as these are owner characteristics that do not vary within owners. The number of observations is greater for the 2- and 3-year survival analyses even though more businesses must be excluded from these (so we can observe complete spells up to two or three years compared to just one – see Tables 9 and 10) because the analyses also are restricted to owners who have different outcomes for their businesses, i.e. they are not all still open, nor all still closed. This eliminates more owners from the 1-year survival sample than from the other two.

These results complement those of Lazear (2005), suggesting that entrepreneurship can be learned. Lazear shows that the number of prior business roles that a person has on the job increases the likelihood of entrepreneurship, but future roles do not.

²² The sample also must be restricted to owners who have different outcomes for their businesses, i.e. they are not all still open, nor all still closed.

Much of the literature on entrepreneurship identifies the traits of entrepreneurs. They are risk takers (Fairlie and Holleran, 2012; Fossen, 2011; Vereshchagina and Hopenhayn, 2009; Kihlstrom and Laffont, 1979), optimists (De Meza and Southey, 1996), wealthy (Blanchflower and Oswald, 1998; Hurst and Lusardi, 2004; Andersen and Nielsen, 2012; Hvide and Moen, 2010, Evans and Jovanovic, 1989, Holtz-Eakin, Joulfaian, and Rosen, 1994), or adept at the recognition of patterns of business formation (Baron and Ensley, 2006). Entrepreneurs tend to group together, suggesting they learn from each other, as when they are physically surrounded by other entrepreneurs (Glaeser, Kerr, and Ponzetto, 2006), or come from a family background of entrepreneurship (Fairlie and Robb, 2007; Lentz and Laband. 1990). But the literature rarely identifies whether these traits are learned or innate.²³ Whereas traits such as risk taking and wealth levels are not very actionable, the evidence here suggests that entrepreneurship skills, even general entrepreneurship skills, per Table 8 results above, can be learned.²⁴

VI. The Success of Large Firms

In all the preceding analyses, we've focused on the survival of the new businesses of small entrepreneurs. That is, we restricted our analyses to the set of businesses of owners who have twenty or fewer establishments. As pointed out in Section III, 95 percent of new establishments are opened by such small owners. However, it may be of interest to see if the patterns of survival for establishments of small firms are mimicked by those of large ones.

They are not. For the establishments of large firms, namely those of owners with 21 or more establishments, there is no gain to experience: The number of prior establishments does not figure into the survival of the current establishment. This is true when we introduce the number of prior establishments in total (columns 1 and 2 of Table 12), or divide the number of prior establishments into those that were opened and remain open versus those that are closed (columns 3 and 4).²⁵

These results are expected. These large firms already have opened at least twenty establishments; in fact, on average, at the time they open a new outlet, these owners have already opened 72.6 establishments in our data (and closed only 14.06 of these). Having a greater number of establishments opened before the current one will not influence the survival of the current one. The larger owners

²³ See Hamilton (2000) for evidence that it is not ability that determines self employment.

²⁴ But see Fairlie et.al (2012) for evidence that an experimental program aimed at teaching entrepreneurship (the GATE program) had some short-run effect, but no long-term effect on the likelihood of business ownership. They also find no evidence of average treatment effects in the short or long run on business performance, household income, or work satisfaction.

²⁵ Interestingly, there is also no duration dependence in these data: the p(Weibull) is valued at 1.00.

have already learned from experience.²⁶ These results are dramatically different from those of small firms where we showed that experience matters, because the vast majority of the latter are owners who open one establishment, close it, and then open a second or third.

Table 12: Weibull Duration Regressions For Establishments of Large Owners

	(1)	(2)	(3)	(4)
Number Opened Before	0.999** (0.000)	1.000 (0.001)		
Squared (Number Opened Before)		1.000* (0.000)		
Number Opened Before Still Open			0.999*** (0.000)	0.999 (0.001)
Sq. (Number Opened Before Still Open)				1.000 (0.000)
Number Opened Before but Closed			1.002 (0.002)	1.007 (0.005)
Sq. (Number Opened Before but Closed)				1.000 (0.000)
Opened in recession	0.815*** (0.062)	0.814*** (0.061)	0.806*** (0.061)	0.807*** (0.061)
Major Chain	0.760*** (0.045)	0.755*** (0.045)	0.768*** (0.045)	0.765*** (0.044)
Urban Establishment	0.911*** (0.032)	0.909*** (0.032)	0.911*** (0.032)	0.911*** (0.032)
Proprietorship	2.262*** (0.304)	2.284*** (0.309)	2.252*** (0.301)	2.254*** (0.301)
Partnership	0.903 (0.068)	0.907 (0.068)	0.909 (0.068)	0.917 (0.069)
Number of obs	120,323	120,323	120,323	120,323
No. of failures	67,580	67,580	67,580	67,580
p (Weibull)	1.01	1.01	1.01	1.02

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Some patterns prevail for the establishments of large firms as well, however. Like those of small owners, establishments of large owners that are opened during a recession or associated with a national brand are much more successful. In fact, the positive effect of opening during recession is even greater for the establishments of large owners than for those of small owners. Being associated with a national brand, which almost half of these establishments are, is not as beneficial to large owners, but being located in an urban environment, which 95 percent of these establishments are, increases their survival, contrary to the effect we found for the establishments of small business owners. Establishments opened by sole proprietors are far less successful overall. As one would

²⁶ See also Benkard, 2000, Levitt, List and Syverson, 2012, and Stith, 2013, for evidence that learning-by-doing occurs almost exclusively early on.

expect, there are also very few such establishments in the set of large owner establishments: less than one percent of these establishments are owned by proprietors, whereas 69 percent of new establishments of small owners were opened by such owners.

The primary difference between large owners and small owners is that establishments opened by large owners survive longer. The mean duration of a new establishment is just over 1200 days, or 3.3 years, for small owners, but almost twice that, at 2100 days, or 5.75 years, for large owners. Therefore, when we look at data on the stock of existing stores, we see evidence that chain stores are a big part of the market. (See Basker et al., 2012, and Jarmin et al., 2009, and our data above.) Because of these much longer durations, large owners operate about 10 percent of all retail establishments at any point in time (compare Tables 1 and 2) even though they open only five percent of the new establishments.

VII. Conclusion

Entrepreneurship is often described as an engine of growth. Entrepreneurs start new businesses, which raise employment and GDP. Small retail entrepreneurs may create fewer jobs per business than large tech entrepreneurs, but small retail is economically important because there are so many establishments. We study all retail businesses in the state of Texas from 1990 to 2011, and during that time there are 2.33 million retail establishments founded in the state by small owners. This includes 46,000 establishments of small owners that are affiliated with national chains via franchising. Large owners (those that have 21 or more establishments in the data period) open only 120,000 establishments, half of which are associated with national brands. The businesses of small owners may be short-lived – the mean duration of these businesses is 3.33 years, with a median of 2 years – but they are prevalent. The establishments in these data span the activities of a broadly defined retail sector – from restaurants, to standard retail stores, to repair shops and hair salons. Among the businesses founded between 1990 and 2011, there are, for example, almost 250 thousand restaurants.

Among small business owners, serial entrepreneurs are more successful. If the owner of a new establishment has owned one prior business, the probability of exit for that new establishment falls by 7 percent. Serial entrepreneurship, however, is not so prevalent in this sector: among our retail establishments of small owners in Texas, 25.5 percent only are owned by those with at least one prior business.

These data show an unexpected pattern underlying serial entrepreneurship for small business owners in retail. The more successful of these owners are those that open a business, close the business, and then open another one. That is, the new business that they open is more likely to survive if their

previous business has been closed than if their previous business remains open. Ownership is sequential. This could be a feature of retail ownership – where most owners are small proprietors and therefore unable to maintain more than one business at a time.

The most noteworthy result is that entrepreneurship is a learned activity. The small business entrepreneur who has opened previous businesses is more successful even when we control for a person fixed effect that would be a proxy for talent. That is, it is not just innate talent that determines future success but past business experience matters as well. These results are consistent with the model and the results of Lazear (2005), who shows that general skills underlie entrepreneurship and a portion of these skills can be learned.

References

- Amaral, A.M., R. Baptista, and F. Lima. 2011. "Serial Entrepreneurship: Impact of Human Capital on Time to Re-Entry." *Small Business Economics*, 37(2011): 1-21.
- Anderson, S., and K. Meisner Nielsen. 2012. "Ability or Finances as Constraints on Entrepreneurship? Evidence from Survival Rates in a National Experiment." *The Review of Financial Studies*, 25(12): 3684-3710.
- Arrow, Kenneth. 1962. "The Economic Implications of Learning by Doing." *The Review of Economic Studies*, 29(3): 155-173.
- Asplund, Marcus and Volker Nocke. 2006. "Firm Turnover in Imperfectly Competitive Markets." *Review of Economic Studies*, 73: 295-327.
- Baron, R.A., and M.D. Ensley. 2006. "Opportunity Recognition as the Detection of Meaningful Patterns: Evidence from Comparisons of Novice and Experienced Entrepreneurs." *Management Science*, 52(9): 1331-1344.
- Basker, Emek, Shawn Klimek and Pham Hoang Van. 2012. "Supersize It: The Growth of Retail Chains and the Rise of the 'Big-Box' Store," *Journal of Economics and Management Strategy*, 21(3): 541-582.
- Bates, T. 1990. "Entrepreneur Human Capital Inputs and Small Business Longevity." *Review of Economics and Statistics*, 72(4): 551-559.
- Benkard, C. Lanier. 2000. "Learning and Forgetting: The Dynamics of Aircraft Production." *The American Economic Review*, 90(4): 1034-1054.
- Birley, S. 1993. "A Comparison of New Businesses Established by 'Novice' and 'Habitual' Founders in Great Britain." *International Small Business Journal*, 12(1): 38-60.
- Blanchflower, D. and A. Oswald. 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics*, 16(1): 26-60.
- Bureau of Labor Statistics, 2010, "Issues in Labor Statistics," Summary 10-09: 1-8
- Cardiff-Hicks, B., F. Lafontaine, and K.L. Shaw. forthcoming. "Do Large Retailers Pay Premium Wages?," *Industrial and Labor Relations Review*.
- Chatterji, A.K. 2009. "Spawned with a Silver Spoon? Entrepreneurial Performance and Innovation in the Medical Device Industry." *Strategic Management Journal*, 30(2): 185-206.
- De Meza, D., and C. Southey. 1996. "The Borrower's Curse: Optimism, Finance and Entrepreneurship." *Economic Journal*, 106(435): 375-386.
- Ennew, C., K. Robbie, and M. Wright. 1997. "Serial Entrepreneurs." *British Journal of Management*, 8(3):251-268.
- Evans, D.S., and B. Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints." *Journal of Political Economy*, 97(4): 808-827.

- Fairlie, R. 2011. "Entrepreneurship, Economic Conditions, and the Great Recession." *IZA Discussion Paper No. 5725*.
- Fairlie, R., and W. Holleran. 2012. "Entrepreneurship Training, Risk Aversion, and Other Personality Traits: Evidence from a Random Experiment." *Journal of Economic Psychology*, 33(2): 366-378.
- Fairlie, R., and A. M. Robb. 2007. "Why are Black-Owned Businesses Less Successful than White-Owned Businesses? The Role of Families, Inheritances, and Business Human Capital." *Journal of Labor Economics*, 25(2): 289-323.
- Fairlie, R. W., Dean Karlan and Jonathan Zinman, 2012. "Behind the GATE Experiment: Evidence on Effects of and Rationales for Subsidized Entrepreneurship Training," NBER Working Papers 17804, National Bureau of Economic Research, Inc.
- Fossen, F. M. 2011. "The Private Equity Premium Puzzle Revisited – New Evidence on the Role of Heterogeneous Risk Attitudes." *Economica*, 78(312): 656-675.
- Fraser, J. A. 1999. Business for sale update. *Inc.* 21(7): 60-69.
- Glaeser, E.L., W.R. Kerr, and G.A.M. Ponzetto. 2006. "Clusters of Entrepreneurship." *Journal of Urban Economics*, 67(2010):150-168.
- Gompers, P., A. Kovner, J. Lerner, and D. Scharfstein. 2005. "Entrepreneurial Spawning: Public Corporations and the Genesis of New Ventures, 1986 to 1990," *Journal of Finance*, 60(2): 517-614.
- Gompers, P., A. Kovner, J. Lerner, and D. Scharfstein. 2006. "Performance Persistence in Entrepreneurship." *Journal of Financial Economics*, 96:18-32.
- Hamilton, B.H. 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment." *Journal of Political Economy*, 108(3): 604-631.
- Holtz-Eakin, D., D. Joulfaian, and H.S. Rosen. 1994. "Entrepreneurial Decisions and Liquidity Constraints." *RAND Journal of Economics*, 25(2): 334-347.
- Hsu, D. 2007. "Experienced Entrepreneurial Founders, Organizational Capital, and Venture Capital Funding." *Research Policy*, 35(5): 722-741.
- Hurst, E. and A. Lusardi. 2004. "Liquidity Constraints, Household Wealth, and Entrepreneurship." *Journal of Political Economy*, 112(2): 319-347.
- Hvide, H. K., and J. Møen. 2010. "Lean and Hungry or Fat and Content? Entrepreneurs' Wealth and Start-Up Performance." *Management Science*, 56(8): 1242-1258.
- Jarmin, R. S., S.D. Klimek, and J. Miranda, 2009. "The Role of Retail Chains: National, Regional and Industry Results," in T. Dunne, J. B. Jensen, and M. J. Roberts (Eds.) *Producer Dynamics: New Evidence from Micro Data*. Chicago, IL: University of Chicago Press.

- Kellogg, Ryan. 2011. "Learning by Drilling: Inter-Firm Learning and Relationship Persistence in the Texas Oilpatch." *Quarterly Journal of Economics*, 126: 1961-2004.
- Kihlstrom, R.E., and J. Laffont. 1979. "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion." *Journal of Political Economy*, 87(4): 719-748.
- Klepper, S. 2001. "Employee Startups in High Tech Industries," *Industrial and Corporate Change*, 10(3):639-674.
- Klepper, S. and S. Sleeper. 2005. "Entry By Spinoffs," *Management Science*, 51(8): 1291-1306.
- Kuechle, G., A.R. Menon, and S.D. Sarasvathy. 2013. "Failing Firms and Successful Entrepreneurs: Serial Entrepreneurship as a Temporal Portfolio." *Small Business Economics*, 40(2): 417-434.
- Lazear, E.P. 2005. "Entrepreneurship." *Journal of Labor Economics*, 23(4): 349-380.
- Lazear, E.P. 2004. "Balanced Skills and Entrepreneurship." *American Economic Review*, 94(2): 208-211.
- Lentz, B.F., and D.N. Laband. 1990. "Entrepreneurial Success and Occupational Inheritance Among Proprietors." *Canadian Journal of Economics*, 23(3): 563-579.
- Levitt, Steven, John List, and Chad Syverson. 2012. "Toward an Understanding of Learning by Doing: Evidence from an Automobile Plant." <http://www.nber.org/papers/w18017>.
- Stith, Sarah. 2013. "Organizational Learning-by-Doing: Evidence from Liver Transplantation," http://sitemaker.umich.edu/sarah.stith/files/stith_olbd_100113.pdf
- Spletzer, James. 2000. "The Contribution of Establishment Births and Deaths to Employment Growth." *Journal of Business and Economic Statistics*, 18(1): 113-126.
- Thompson, Peter. 2012. "The Relationship between Unit Cost and Cumulative Quantity and the Evidence for Learning-by-Doing." *Journal of Economic Perspectives*, 26(3): 230-224.
- Vereshchagina, G., and H. A. Hopenhayn. 2009. "Risk Taking by Entrepreneurs." *The American Economic Review*, 99(5): 1808-1830.
- Westhead, P., and M. Wright. 1998. "Novice, Portfolio, and Serial Founders: Are They Different?" *Journal of Business Venturing*, 13(3): 173-204.
- Westhead, P., M. Wright, D. Ucbasaran. 2009. "The Extent and Nature of Opportunity Identification by Experienced Entrepreneurs." *Journal of Business Venturing*, 24(2): 99-115.
- Zhang, J. 2011. "The Advantage of Experienced Start-Up Founders in Venture Capital Acquisition: Evidence from Serial Entrepreneurs." *Small Business Economics*, 36(2): 207-228.
- Zimmerman, Martin B. (1982). "Learning Effects and the Commercialization of New Energy Technologies: The Case of Nuclear Power." *The Bell Journal of Economics*, 13(2): 297-310

Appendix A: Data Appendix

Our data set of retail and service businesses in Texas is uniquely comprehensive. We can identify every entity that collected sales taxes from consumers in the state—and hence every establishment that sold goods or services to end consumers—between 1990 and 2011. We constructed this database by combining information from several downloads of the Texas Sales and Use Tax Permit Holder Information File. After grouping records relating to the same business establishment, eliminating a small set of non-business entities (mostly associations that have some sales activities, as well as some public entities such as school districts), some with missing data (especially on end-of-business dates), and those that are not opened for at least 30 days (on the presumption that these are not real businesses), and 18 observations with missing data on urban status (defined as whether they are established within city limits in the Texas Sales Tax files), we obtain a sample size of 3,200,824 businesses owned by 2,160,391 separate owners. We use this full set of observations when determining owner size, which is measured in terms of number of businesses. Of these businesses, 2,818,505, owned by 1,956,829 separate owners, are started after Jan. 1, 1990 and thus are used in the calculation of number of businesses “established before” the businesses of interest in our analyses.

For our analyses, we reduce the sample to the businesses that indicate that they operate in the retail sector, where we define retail broadly given that the presence of these businesses in the data already indicates that they collect sales taxes.²⁷ We have 2,780,370 such businesses, owned by 1,890,321 owners, in our original sample of 3,200,824 businesses. Of these, 2,452,311 businesses, owned by 1,715,352 owners, are started from 1990 onward (see also Tables 1 and 2).

Our interest in small business entrepreneurship leads us to focus most of our analyses on those owners with 20 or fewer businesses. Our sample of retail businesses started from 1990 onward and owned by small owners is 2,331,988 businesses for 1,713,112 owners. In other words, we have only 2240 large owners, and these are associated with only 120,323 retail establishments founded since Jan. 1, 1990. Because these establishments have longer duration, they represent about 10 percent of the total number of retail establishments operating at a given point in time (again see Tables 1 and 2).

²⁷ Specifically, we include wholesalers as well as all retail trade, and food services and entertainment and all personal services in our definition of retail. See Table A1 below for a list of 3-digit NAICS industries we include.

Table A1: The Distribution of Businesses by 3-digit NAICS for our Retail Sample

NAICS 2002 Sector	Frequency	Percent	Cumulative
423: Merchant Wholesalers Durable Good	114,599	4.12	4.12
424: Merchant Wholesalers Nondurable Good	49,950	1.80	5.92
425: Wholesale Electronic Markets and A	4,667	0.17	6.09
441: Motor Vehicle and Parts Dealers	97,016	3.49	9.58
442: Furniture and Home Furnishings Stores	59,022	2.12	11.70
443: Electronics and Appliance Stores	80,483	2.89	14.59
444: Building Material and Garden Equipment	40,417	1.45	16.05
445: Food and Beverage Stores	67,181	2.42	18.46
446: Health and Personal Care Stores	48,002	1.73	20.19
447: Gasoline Stations	69,925	2.51	22.70
448: Clothing and Clothing Accessories	215,140	7.74	30.44
451: Sporting Goods Hobby Book and Misc.	132,346	4.76	35.20
452: General Merchandise Stores	63,145	2.27	37.47
453: Miscellaneous Store Retailers	548,070	19.71	57.19
454: Nonstore Retailers	101,968	3.67	60.85
531: Real Estate	9,672	0.35	61.20
532: Rental and Leasing Services	39,961	1.44	62.64
533: Lessors of Nonfinancial Intangible Assets	128	0.00	62.64
541: Professional Scientific and Technical Serv.	195,188	7.02	69.66
551: Management of Companies and Enterprise	1,351	0.05	69.71
561: Administrative and Support Service	203,687	7.33	77.04
562: Waste Management and Remediation Serv.	3,638	0.13	77.17
711: Performing Arts Spectator Sports	16,972	0.61	77.78
712: Museums Historical Sites	998	0.04	77.81
713: Amusement Gambling, and Recreation	24,152	0.87	78.68
721: Accommodation	8,916	0.32	79.00
722: Food Services and Drinking Places	303,126	10.90	89.91
811: Repair and Maintenance	180,234	6.48	96.39
812: Personal and Laundry Services	100,416	3.61	100.00
Total	2,780,370	100.00	

Appendix B: Robustness Analysis

Table A2: Weibull Duration Regressions, with Industry (3-digit NAICS) Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Number Opened Before	0.978*** (0.001)	0.952*** (0.002)			0.956*** (0.003)	0.916*** (0.003)
Squared (Number Opened Before)		1.004*** (0.000)				1.009*** (0.001)
Number Opened Before Still Open			0.997 (0.003)	1.002 (0.004)		
Sq. (Number Opened Before Still Open)				0.999* (0.001)		
Number Opened Before but Closed			0.963*** (0.001)	0.927*** (0.002)		
Sq. (Number Opened Before but Closed)				1.011*** (0.000)		
Opened in recession	0.927*** (0.003)	0.927*** (0.003)	0.927*** (0.003)	0.927*** (0.003)	0.931*** (0.003)	0.928*** (0.003)
Major Chain	0.697*** (0.007)	0.696*** (0.007)	0.691*** (0.007)	0.690*** (0.007)	0.689*** (0.006)	0.692*** (0.007)
Urban Establishment	1.203*** (0.003)	1.203*** (0.003)	1.202*** (0.003)	1.202*** (0.003)	1.192*** (0.003)	1.190*** (0.003)
Proprietorship	1.621*** (0.004)	1.624*** (0.004)	1.628*** (0.004)	1.631*** (0.004)	1.621*** (0.004)	1.624*** (0.004)
Partnership	1.717*** (0.007)	1.715*** (0.007)	1.719*** (0.007)	1.719*** (0.007)	1.717*** (0.007)	1.716*** (0.007)
Number Opened Before * Recession					0.991*** (0.003)	1.002 (0.005)
Number Opened Before * Chain					1.012** (0.006)	1.016 (0.011)
Number Opened Before * Urban					1.025*** (0.003)	1.044*** (0.004)
Sq. (Number Opened Before * Recession)						0.998*** (0.001)
Sq. (Number Opened Before * Chain)						0.998* (0.001)
Sq. (Number Opened Before * Urban)						0.995*** (0.001)
NAICS Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs	2,331,988	2,331,988	2,331,988	2,331,988	2,331,988	2,331,988
No. of failures	1,849,592	1,849,592	1,849,592	1,849,592	1,849,592	1,849,592
p (Weibull)	0.89	0.89	0.89	0.89	0.89	0.89

Notes: Standard errors, clustered at the owner level, in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$