THE IMPACT OF REGIONAL AGE STRUCTURE ON ENTREPRENEURSHIP IN CANADA

by

Liang Ma

Bachelor of Economics, Shandong University of Finance and Economics, 2015

A Report Submitted in Partial Fulfilment of the Requirements for the Degree of

Master of Arts

in the Graduate Academic Unit of Economics

Supervisor: Weiqiu Yu, PhD, Dept. of Economics

Examining Board: Ted McDonald, PhD, Dept. of Economics

Philip Leonard, PhD, Dept. of Economics

This report is accepted by the Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

July, 2016

Abstract

Due to decreasing birth rate and mortality rate in many countries, ageing has become a global issue in today's world. Ageing directly impacts business and the overall economy of a nation. Empirical studies have found an inverse U-shaped relationship between age and the decision to start a business using micro-level data. Some studies have shown that the decision to become an entrepreneur is a regional phenomenon. Bronte, Falck & Heblich (2009) explored the effect of demographic factors and age on business startups in Germany using an aggregate data set and a count data model. They found that an inverse U-shaped relationship between the regional age structure and start-up activity in a region at the aggregate level. This report adopts a similar methodology and uses a longitudinal data set to examine the relationship between age structure and regional business activities in Canada from 1988 to 2014. Results show that differences in the age structures do contribute to the variation in the business activities across the Canadian provinces, which has strong policy implications for provincial governments to design policies directed at promoting entrepreneurial activities within their provinces. Moreover, our findings suggest that the age-specific likelihood of becoming an entrepreneur changes over time, indicating the existence of age-specific peer effects.

Acknowledgement

I would like to thank my supervisor, Dr. Weiqiu Yu, for her guidance, suggestions, supervision and support in the preparation of this report. I would also like to thank the other members of my committee, Dr. Ted McDonald and Dr. Philip Leonard for their useful comments as readers.

Thanks are also extended to my family members and friends. Without their support, I could not complete my MA degree.

Table of Contents

ABSTRACT II

ACKNOWLEDGEMENT III

I. INTRODUCTION 1

II. LITERATURE REVIEW 5

III. METHODOLOGY AND DATA 11

3.1 METHODOLOGY 11

3.2 DATA 13

IV. RESULTS AND DISCUSSION 27

V. CONCLUSIONS 30

VI. REFERENCES 35

APPENDIX 38

1.1 POISSON REGRESSION MODEL 38

1.2 FIXED-EFFECTS MODEL 38

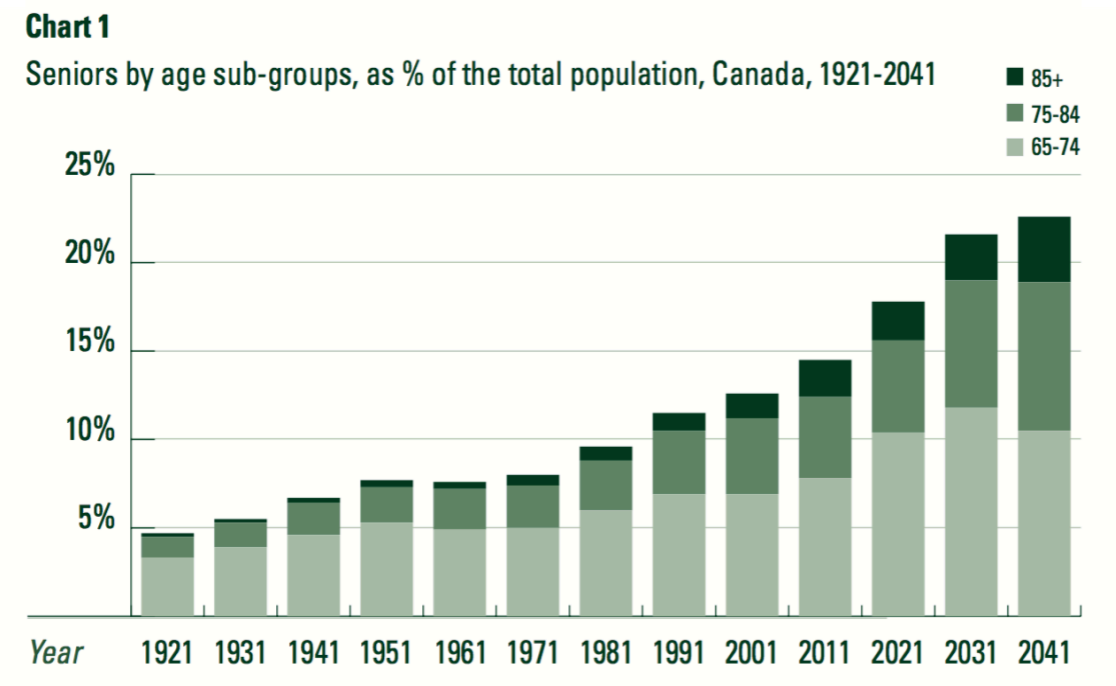
1.3 POISSON REGRESSION MODEL VS OLS REGRESSION MODEL 39

2.1 ADDITIONAL REGRESSION RESULTS 39

CURRICULUM VITAE

I. Introduction

A sustained decline in mortality and fertility rates during the twentieth century has resulted in a shift towards older populations worldwide. In Canada as shown in Chart 1, the number of seniors in all age groups is expected to continue to rise, and by 2041, seniors are projected to comprise nearly a quarter (24.5%) of the Canadian population, as compared to 14.8% today (Health Canada). The first cohort of the baby boomers turned 65 in 2011, and the last baby boom cohort will turn 65 in 2029.



Source: Health Canada.

Furthermore, Canada's senior population is distributed unevenly across the provinces, with the highest concentration in the Atlantic Provinces. Nova Scotia has the highest proportion of seniors, at 16.6% of its population, followed by New Brunswick (16.5%) and Prince Edward Island (16.3%). Demographics is one of the most important factors affecting entrepreneurship, job creation and innovation.

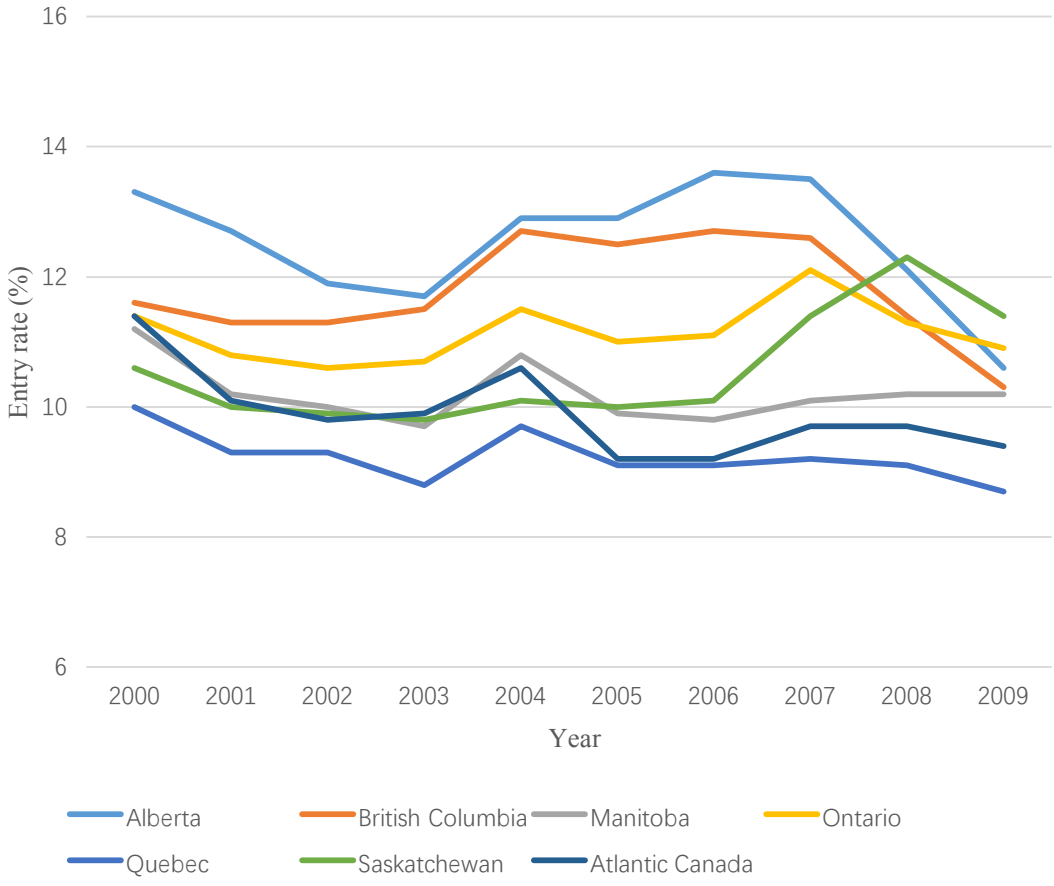
A great deal of research has argued that entrepreneurs are the lifeblood of the economy. For one thing, from the perspective of the economy, the development of entrepreneurial activities could promote the adjustment and transference of economy structure. In the process, more employment opportunities may be generated and there will be an increase in employment rate, hence stable economic development. For another, as individuals, some people regard to be an entrepreneur as a pursuit due to a variety of motives: unemployment, job insecurity, a burning idea, a desire for freedom, a wish to make a difference, and economic opportunities. Both show the significance of entrepreneurship. However, different age groups have distinct features and characteristics to starting a new business. For example, people in their twenties are considered the embodiment of great passion and courage when having a try on a new business, but little experience at the same time, while people aged forty and above have a better command of the knowledge, social networking as well as their experience and abilities to solve problems but less chance to fail. In the meantime, the emergence of senior entrepreneurship has strong policy implications for an ageing society.

To understand the relationship between regional ageing structure and entrepreneurial activities in Germany, Bronte, Falck & Heblich (2009) explored the effect of demographic factors and age on business start-ups in Germany. Using an aggregate data set and a count data model, they found there exists an inverse U-shaped relationship between the regional age structure and start-up activity in a region, which may be, as they stated, due either to individual age effects or to age-specific

peer effects. The latter, which involves in the entrepreneur's social networking, could make the peers get access to more resources and information about the opportunities and risks, hence the reduction in costs and failure. In addition, age could affect the peer's social networking, namely that people incline to make friends with those of a similar age, and thus, they possess more experience and social networking and when people perceive their peer's behaviors, they will be more motivated to make their own decisions to start a new business as well. Therefore, as Zhang and others have discussed previously, learning more about the age-specific peer effects that support entrepreneurial activity is a valuable contribution to the newly emerging field of study that is investigating the incentives for entrepreneurship in the context of an ageing society (Zhang (2008); Bronte, Falck & Heblich (2009)).

Since there are considerable differences across the Canadian provinces in both age structures and entrepreneurial activities, this study is highly relevant to Canada. With regards to entrepreneurial activities, Figure1 shows that the entry rates in six Canadian provinces and the Atlantic region. In particular, while British Columbia, Alberta, Saskatchewan and Ontario were consistently higher than the average, Atlantic Canada and Quebec experienced the lowest and declining entry rates from 2000 to 2009.

Figure 1
Annual Entry Rates of Firms in Canada, 2000-2009



Source: Baldwin, Liu & Wang (2013).

To examine the relationship between age structure and business activities in Canada, we adopt a similar methodology used by Bronte, Falck & Heblich (2009) to analyze a longitudinal data set consisting of ten Canadian provinces from 1988 to 2014. Results show that differences in the age structures do contribute to the variation in the business activities of ten Canadian provinces, which has strong policy implications for regional governments to design policies directed to increasing entrepreneurial activities within their regions. Moreover, we also find the peak of age cohorts starting a business moves from the older groups to the younger generations,

which sheds clear light on the impact of the age-specific peer effects.

The rest of the report is organized as follows: Section II contains the literature review; Section III explains methodology and data used in this study; Section IV presents and discusses the results; and Section V concludes.

II. Literature Review

Many theoretical and empirical studies have been done to emphasize the importance of entrepreneurship in driving economic growth (e.g., Schumpeter (1951)). Wang, Chen & Zhu (2014) define entrepreneurship as an ability that an entrepreneur who has to establish and manage his or her own business. It is an important and intangible factor of production. Audretsch and Feldman (1996) found that entrepreneurship is an important source of technological development and innovation. In general, the effect of entrepreneurship promoting economic growth can be divided into micro and macro levels. At the micro level, entrepreneurship can promote establishments of private enterprises, thereby creating employment. At the macro level, the new growth theory since the 1990s states that entrepreneurship is a shared system of values of entrepreneurs, embodied in the specific nature such as innovation, anti-risk, work attitude and their own sense of worth. These traits can affect the innovation behavior of entrepreneurs, and are important factors in determining total factor productivity and economic growth.

Li et al. (2009) provided evidence to support the importance of entrepreneurship empirically. Their study examined the impact of entrepreneurship on economic

growth by using a panel data set of 29 provinces in China over 20 years. By adopting the system generalized method of moments to estimate the growth regression framework including indicators of entrepreneurship, they found that entrepreneurship has a significantly positive effect on economic growth. This finding is consistent with existing findings in the literature[1]. Therefore, it is essential to study the influencing factors of the entrepreneurship.

In the literature on the determinants of entrepreneurship, various factors including demography, age, individual characteristics, peer effects, and surrounding environment have been identified to affect the decision of individual to start a business. Some studies attempt to use self-reported data with Logit model to explain the probability of becoming an entrepreneur[2]. Typical explanatory variables include age, gender, race, education, earnings, capital assets, previous professional experience, previous employment status (i.e., wage, part-time, full-time), marital status, professional status of the parents, and scores from psychological tests. Their findings suggest that age could affect individual's occupational choice.

Stangler and Spulber (2013) investigated the effect of age-related demographic changes on entrepreneurship in the United States. They found that ageing causes the decline in entrepreneurship. However, this does not mean that there is absence of entrepreneurship at old age since the labour force remains stable and business continues to grow in spite of demographic changes. Demographic analysis has an

[1] See Wong, Ho & Autio (2005), Carree and Thurik (2010).
[2] See Parker (2004), Blanchflower (2004) and Grilo and Thurik (2008).

advantage of identifying population trends that may affect individual decision making in different times. This means today's children influence the world of the next two or three decades. However, because so many factors could affect an individual's decision to become an entrepreneur, there can be no definite connection between demographic changes and aggregate entrepreneurship.

Levesque and Minniti (2006) investigated the role of age in motivating entrepreneurial behavior. They found that there exists a negative relationship between age and entrepreneurial activities. The paper builds upon the Becker's theory of time allocation (Becker, 1965) to construct a model that individual's motivations for selection of a particular career path according to the dynamic interplay of both contextual factors, as features of the socio-economic environment within which entrepreneurial decisions are taken and inherent factors, as characteristics of the individual. Recent research implies that entrepreneurial behavior is an embedded phenomenon. From a macro standpoint, the inducements to entrepreneurs to start a new business are largely influenced by socio-economic environment and could vary dramatically in different districts. For any individual, entrepreneurial environment cannot be existed without binding constraints from socio-economic contextual circumstances. From the perspective of the individuals, however, entrepreneurial decisions are not always relying on environment factors. Their own inherent characteristics may be a contributing factor as well.

The authors suggest that the effect of age on entrepreneurial decisions is analogous to that of an inherent factor. Therefore, they looked at the interdependency of both the

contextual and the inherent factors on entrepreneurial behavior. They argued that with ageing process an individual is gaining more experience and seniority in waged labour hence the decline of their incentives to starting new firms (or increasing the opportunity cost of starting a new firm). This study has significant policy implications as it suggests that things need to change along with ageing populations or else there might be a decline in entrepreneurship which may in turn decrease economic growth.

Kurek and Rachwal (2011) examined the relationship between ageing population and development of entrepreneurship within the European Union (EU). Like in Canada, both longer life expectancy and declining fertility rates cause the EU faces a similar demographic challenge of a shrinking labour force and rising dependency ratio. The authors suggest that population ageing will have an impact on the development of entrepreneurship. To deal with this challenge, European countries are increasingly seeking ways to identify the knowledge, skills and attributes that will allow their senior citizens to play an active part in the dynamic society. The study concludes that in addition to demographics, local economic policies, legal and administrative factors, social security systems, cultural factors, and market all have influence over development of entrepreneurship.

Webber and Schaper (2003) provided an overview of research into the phenomena of the so-called “grey entrepreneur” based on the research from Australia, the United Kingdom, New Zealand and the USA. In simple words, grey entrepreneurs are

entrepreneurs who are elderly and own a business of their own. They are also known as senior entrepreneurs, older entrepreneurs and third-age entrepreneurs. Long term demographic trends throughout the developed world suggest that such entrepreneurs will come to play an increasingly important part of economic activity, as populations age and the traditional workforce age cohort declines. The paper discusses various advantages and disadvantages of grey entrepreneurship in an ageing society. The authors state "in Australia, almost a third of small business owners are over 50 years of age, and their numbers are growing faster than in many younger age cohorts (Australian Bureau of Statistics 2001b)" and "almost more than 40% of the self-employed persons are aged 50 and above in New Zealand (Statistics New Zealand 2001)" which support the "grey entrepreneur" (reason being decreasing birth rates and increasing life expectancy) and its impact on the workforce is evident. The authors point out that a number of potential factors appear to have an impact on the activities and likely success of grey entrepreneurs. Characteristics of grey entrepreneurs include gender, high levels of autonomy, independence, self-reliance, high levels of personal fitness and strong levels of family support. The experience they possess, better networking, greater financial resources are the advantages of the grey entrepreneurship.

Finally, Bronte, Falck & Heblich (2009) studied the impact of regional age structure on entrepreneurship. Previous empirical studies using individual data set have found that there is strong evidence for an inverse U-shaped relationship between age and start up activity. Other studies found that becoming an entrepreneur is a regional

event, with potential entrepreneurs benefiting from their local networks. This paper links both strands of literature by introducing age-specific peer effects. The authors argue "at the regional level, the impact of a region's age structure on regional start-up activity should not change over time if only the individual age-specific effect is prevalent; however, the impact of a region's age structure on start-up activity should change over time and with the ageing of the population if age-specific peer effects are prevalent." To explore the relationship between the regional age structure and start-up activity in a region, they used an aggregate data set containing the age distribution of the population of western German regions over time and a count-data method found that, consistent with empirical studies using individual data set, an inverse U-shaped relationship exists between the regional age structure and start-up activity at the aggregate level. With this finding, one can determine changes in the entrepreneurial activities of a region using changes in the age structure of that region and design age appropriate policies to enhance regional economic performance.

Based upon the literature review, we can see that entrepreneurship plays an important role in economic growth. And many researchers adopted different methods to study the determinants of entrepreneurship at micro level. Empirical finding in Bronte, Falck & Heblich (2009) is highly relevant to Canada where population is not only ageing rapidly but age structure differs significantly across the Canadian provinces. Accordingly, we adopt a similar approach to analyzing an aggregate data set for Canada and examining the relationship between provincial age structure and entrepreneurial activities in Canada.

III. Methodology and Data

3.1 Methodology

Following Bronte, Falck & Heblich (2009), we use a count data model to examine the impact of provincial age structure on the aggregate entrepreneurial activities in each of the ten Canadian provinces from 1988 to 2014. While Bronte, Falck & Heblich (2009) used the number of start-ups, we use the total number of establishments as the dependent variable in order to capture the aggregate level of business activities in each province. The reason is: Statistics Canada only provides the information of start-ups (number of entrants) counts from 2009 to 2013. Time series (with-in) variations from that are not enough to perform our regression. Similar to that in Bronte, Falck & Heblich (2009), the unit of measurement is the establishment, not the firm which includes both firm headquarters and subsidiaries.

The estimated equation is:

$$Establishments_{it} = f(Population_{it}, AgeStructure_{it}, UnemploymentRate_{it}) \quad (1)$$

Where: $Establishments_{it}$ = total establishments in province i and year t,

$Population_{it}$ = population of province i in year t, used to control for the size of province i,

$AgeStructure_{it}$ = the age structure of province i in year t and

$UnemploymentRate_{it}$ is used to control for the general economic condition of province i in year t.

Since the number of Establishments $_{it}$ is a count variable, using classical linear regression model to estimate Equation (1) is not appropriate. Therefore, the following Poisson model of regression[3] is used to estimate the relationship between the total establishment and age structure of a province in Canada.

$$E\,(Establishments_{it} \mid Population_{it}, AgeStructure_{it}, UnemploymentRate_{it}) = \lambda_{it}$$

$$= \exp(\alpha_i + \beta_1 Population_{it} + \beta_2 AgeStructure_{it} + \beta_3 UnemploymentRate_{it}) \quad (2)$$

where λ_{it} is the expected number of establishments in province i and year t; α_i is added to capture most of the time-invariant and unobserved provincial characteristics, including various policies that are aimed at promoting local business. By using fixed-effects[4], even if the unobserved regional characteristics are correlated with our explanatory variables, we could still get consistent estimates.

In Poisson models, there is always a risk that the observed variance is greater than the mean, which leads to overdispersion. To control for this problem, we use robust standard errors as described by Wooldridge (1999). Also according to Chen (2013), even the data are not Poisson distributed, the estimation results will still be robust against misspecifications. However, we also ran negative binomial regressions for robustness check and to control for potential overdispersion.

Like in Bronte, Falck & Heblich (2009), we split the sample period from 1988 to 2014 into two periods (1988-2000 and 2001-2014) and run separate regressions to examine the structural stability of the relationship between age structure and

[3] See Appendix 1.1.
[4] See Appendix 1.2.

aggregate entrepreneurial activity. As well we break the sample in 1997 for checking the robustness of our results.

3.2 Data

The data used in this study are all from Statistics Canada. Total establishments are extracted from the Canadian Business Patterns (CBP) database for ten Canadian provinces from 1988 to 2014. Canadian business patterns data provide counts of active locations by industry classification and employment size for Canada and the provinces. The data are compiled from the Business Register, which is a repository of information on the Canadian business population. The population data by age group: 20-29, 30-39, 40-49, and 50-64 are extracted from Statistics Canada's key socioeconomic database (CANSIM) Table 051-0001 - Estimates of population, by age group and sex by year and by province. And the annual average unemployment rates are from Statistics Canada: Table 282-0086 - Labour force survey estimates (LFS), supplementary unemployment rates by province CANSIM (database).

Table 1 presents the summary statistics of all the variables used in this study. It is easy to find that the standard deviation of Establishments (i.e., our dependent variable) is greater than its mean value, indicating that there is a problem of overdispersion. Looking at the age structure reported as the population shares of the four different age groups reveals the extent of variation across the provinces and ageing population for all provinces over time. Furthermore, our sample periods from 1988-2014 cover baby boomers' (1946~1964) early years (24~42) and later years

(50~68). Figure 2 demonstrates the growth rate of total establishments in ten different provinces from 1989 to 2014. We could see that the growth rate in different provinces reached its peak in 1997-1998 simultaneously, where the baby boomer cohorts are at their forties. However, we should notice that the growth rate of establishments for that time period is too high to believe. For robustness check, we also split our data from 1997(1988-1997 and 1998-2014) and ran regressions separately.

Table 1: Summary Statistics

	1988-2014				1988-2000				2001-2014			
Variable	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Establishments (Counts)	179419.5	225557.7	6120	999766	115903.5	139921	6120	701915	238398.6	270208	10225	999766
Population Size (in 10,000)	310.1901	364.0663	12.9289	1367.874	288.231	333.109	12.9289	1168.329	330.5807	390.7148	13.6665	1367.874
Share of Population 20-29 (%)	14.19837	1.563789	11.66	19.26	15.10062	1.546964	12.77	19.26	13.36057	1.021424	11.66	16.24
Share of Population 30-39 (%)	15.06537	1.930703	11.53	19.14	16.72215	1.011145	13.92	19.14	13.52693	1.14629	11.53	16.32
Share of Population 40-49 (%)	14.81619	1.488745	10.15	17.07	14.12277	1.502936	10.15	16.68	15.46007	1.151364	12.07	17.07
Share of Population 50-64 (%)	16.76674	3.494227	11.27	23.96	13.74254	1.31304	11.27	17.09	19.57493	2.352099	14.23	23.96
Unemployment Rate (%)	9.245185	3.620434	3.2	20.1	10.54308	3.744307	5	20.1	8.04	3.052344	3.2	16.5

Figure 2: Growth Rate of Toal Establishments by province, 1989-2014

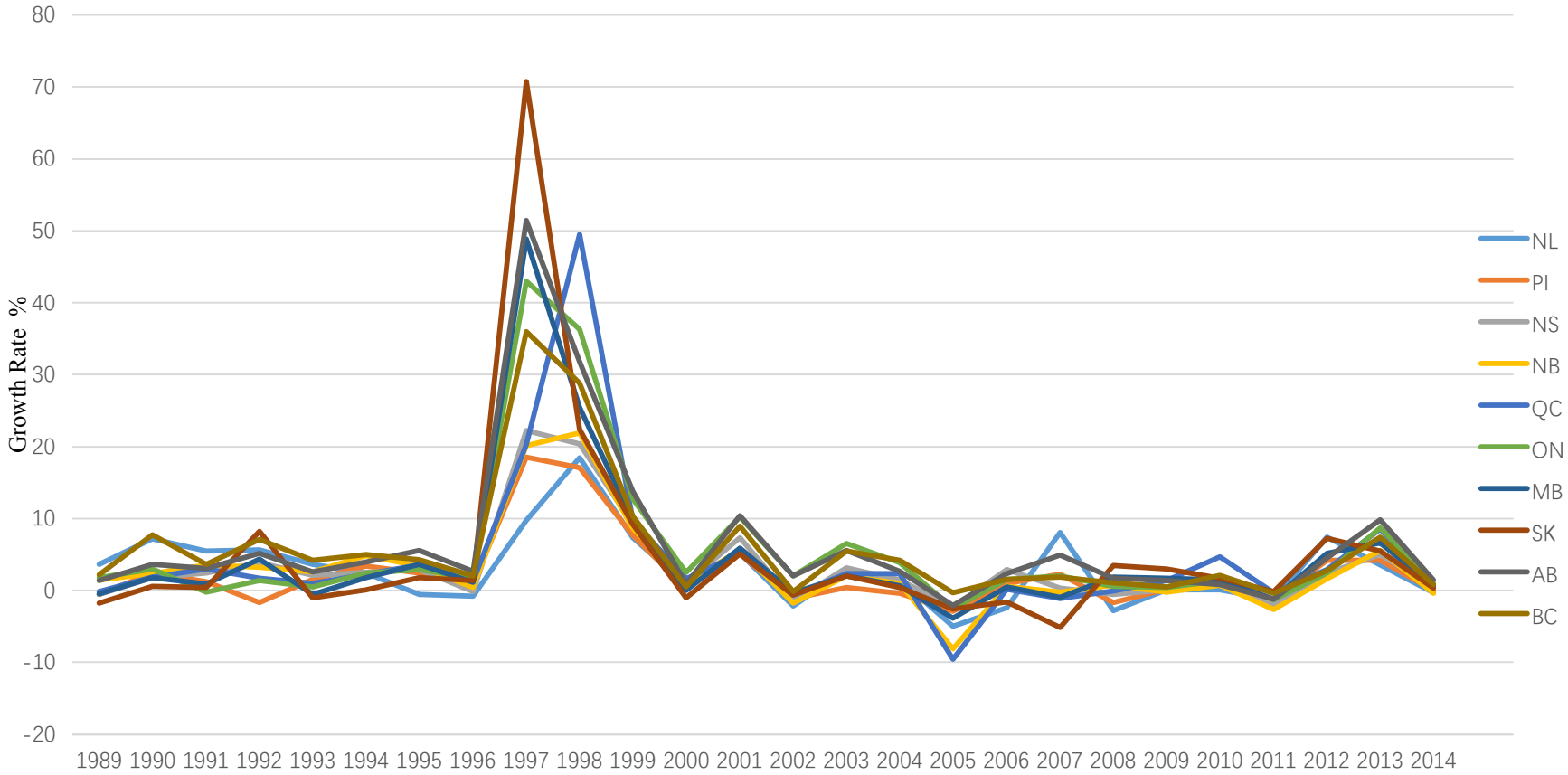


Table 2: Regression Results

Results, 1988-2014						
Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0025785***	0.000369	0.000	0.0033841***	0.0006913	0.001
Share of Population 20-29 (%)	0.0045585	0.0261975	0.862	0.0457045	0.0325689	0.194
Share of Population 30-39 (%)	0.0724852**	0.0289535	0.012	-0.0325727	0.0365149	0.396
Share of Population 40-49 (%)	**0.1246396*** **	0.023227	0.000	**0.1299821*** **	0.0253994	0.001
Share of Population 50-64 (%)	0.0880996***	0.0233283	0.000	0.0326462	0.0223996	0.179
Unemployment Rate (%)	-0.0670746***	0.0121085	0.000	-0.0432223**	0.014028	0.013
Number of Observations	270			270		
Wald chi2/F test	3322.39***		0.000	1538.4***		0.000
Test of Equality of "population share" coefficients F test	474.1***		0.000	251.3***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	631.32***		0.000	600.55***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 2: continue

Results, 1988-2014

Method	Panel C: Dependent Variable=Establishments FE NB Regression with Cluster (Province)			Panel D: Dependent Variable=Establishments per capita FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.001597***	0.0001948	0.001	0.0001076***	0.0000274	0.003
Share of Population 20-29 (%)	-0.0326306	0.0149249	0.194	0.0037239*	0.0018701	0.078
Share of Population 30-39 (%)	-0.0583433	0.0234181	0.396	-0.0025003	0.0014801	0.125
Share of Population 40-49 (%)	**0.0804535*** **	0.012775	0.001	**0.007195*** **	0.0017138	0.002
Share of Population 50-64 (%)	0.021337	0.0126615	0.179	0.0014131	0.0009733	0.180
Unemployment Rate (%)	-0.0343928**	0.0104286	0.013	-0.0020047***	0.0005198	0.004
Number of Observations	270			270		
Wald chi2/F test	1494.41***		0.000	281.3***		0.000
Test of Equality of "population share" coefficients F test	238.86***		0.000	274.39***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	367.82***		0.000	206.54***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 3: Regression Results

Results, 1988-2000

Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0043583***	0.0003541	0.000	0.0034885***	0.0005017	0.000
Share of Population 20-29 (%)	0.02247	0.0215959	0.298	-0.0053847	0.029276	0.858
Share of Population 30-39 (%)	-0.2366942***	0.0353813	0.000	-0.3026752***	0.0460399	0.000
Share of Population 40-49 (%)	**0.1063112*** **	0.0307534	0.001	**0.1069953*** **	0.0231336	0.001
Share of Population 50-64 (%)	0.0221319	0.0252406	0.381	-0.0187098	0.020939	0.395
Unemployment Rate (%)	-0.0201446**	0.0086517	0.020	-0.0063326	0.0101615	0.549
Number of Observations	130			130		
Wald chi2/F test	71195.9***		0.000	343.05***		0.000
Test of Equality of "population share" coefficients F test	163.42***		0.000	43.98***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	372.87***		0.000	137.77***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 3: continue

Results, 1988-2000

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.0000837	0.0004668	0.858	0.000087**	0.0000375	0.046
Share of Population 20-29 (%)	-0.0759615***	0.0215638	0.000	0.0000626	0.0018616	0.974
Share of Population 30-39 (%)	-0.3905766***	0.0330669	0.000	-0.0180177***	0.0033506	0.000
Share of Population 40-49 (%)	**0.0764922*** **	0.0230331	0.001	**0.0052451** **	0.0018567	0.020
Share of Population 50-64 (%)	-0.0634452***	0.0161794	0.000	-0.0021549	0.0013423	0.143
Unemployment Rate (%)	0.0082719	0.0105992	0.435	0.0001965	0.0005675	0.737
Number of Observations	130			130		
Wald chi2/F test	1131.3***		0.000	85.45***		0.000
Test of Equality of "population share" coefficients F test	507.13***		0.000	25.35***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	720.1***		0.000	60.64***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 4: Regression Results

Results, 2001-2014						
Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0013071***	0.0003529	0.000	0.0011627***	0.0003305	0.007
Share of Population 20-29 (%)	**0.0516854***	0.0301942	0.087	**0.0316696**	0.0196329	0.141
Share of Population 30-39 (%)	0.0299517	0.0259183	0.248	0.0002969	0.0185081	0.988
Share of Population 40-49 (%)	-0.0048123	0.0200023	0.810	-0.0307353**	0.0117031	0.028
Share of Population 50-64 (%)	0.010848	0.0172239	0.529	-0.0070923	0.0123067	0.579
Unemployment Rate (%)	-0.0077996**	0.0038984	0.045	-0.008308*	0.0043034	0.086
Number of Observations	140			140		
Wald chi2/F test	5194.73***		0.000	12.76***		0
Test of Equality of "population share" coefficients F test	22.04***		0.000	21.93***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	22.08***		0.000	16.66***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 4: continue

Results, 2001-2014

Method	Panel C: Dependent Variable=Establishments FE NB Regression with Cluster (Province)			Panel D: Dependent Variable=Establishments per capita FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.0007975***	0.0001506	0.000	0.0000203	0.0000152	0.215
Share of Population 20-29 (%)	**0.0369001*** **	0.0103389	0.000	**0.0007974**	0.0010303	0.459
Share of Population 30-39 (%)	0.003406	0.0100045	0.734	-0.0020438**	0.0007161	0.019
Share of Population 40-49 (%)	-0.0273234***	0.0082314	0.001	-0.0021787***	0.0004968	0.002
Share of Population 50-64 (%)	-0.0011735	0.0072527	0.871	-0.0019629***	0.000465	0.002
Unemployment Rate (%)	-0.0045215	0.0036845	0.220	-0.0004662*	0.0002366	0.080
Number of Observations	140			140		
Wald chi2/F test	251.3***		0.000	6.53***		0.000
Test of Equality of "population share" coefficients F test	56.64***		0.000	3.8***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	63.68***		0.000	5.16***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 5: Regression Results

Results, 1988-1997

Method	**Panel A: Dependent Variable=Establishments** **FE Poisson Regression with Cluster (Province)**			**Panel B: Dependent Variable=log(Establishments)** **FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0030783***	0.0004548	0.000	0.0023738***	0.0003559	0.000
Share of Population 20-29 (%)	0.0415538**	0.0187425	0.027	0.0076785	0.0172196	0.666
Share of Population 30-39 (%)	-0.1897468***	0.0438706	0.000	-0.2020543***	0.0381199	0.000
Share of Population 40-49 (%)	**0.1221206*** **	0.0221585	0.000	**0.1041954*** **	0.0162145	0.000
Share of Population 50-64 (%)	0.0311094	0.0464496	0.503	-0.0130089	0.0266267	0.637
Unemployment Rate (%)	-0.0073206***	0.0021131	0.001	-0.0031267	0.0048569	0.536
Number of Observations	100			100		
Wald chi2/F test	85591.4***		0.000	150.38***		0.000
Test of Equality of "population share" coefficients F test	139.39***		0.000	53.47***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	322.26***		0.000	105.58***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 5: continue

Results, 1988-1997

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.0010494***	0.0003827	0.006	0.0000327*	0.0000173	0.091
Share of Population 20-29 (%)	-0.008194	0.0168747	0.627	-0.0001636	0.0008901	0.858
Share of Population 30-39 (%)	-0.2428398***	0.0436146	0.000	-0.0101447***	0.0021937	0.001
Share of Population 40-49 (%)	**0.1073664*** **	0.0211868	0.000	**0.0036013*** **	0.0009233	0.004
Share of Population 50-64 (%)	-0.0336201	0.0257389	0.191	-0.0012403	0.0010518	0.269
Unemployment Rate (%)	0.003162	0.0077286	0.682	0.0000538	0.0002206	0.813
Number of Observations	100			100		
Wald chi2/F test	258.91***		0.000	48.25***		0.000
Test of Equality of "population share" coefficients F test	158.39***		0.000	17.18***		0.001
Test of "population share" coefficients simultaneously equal to zero F test	222.15***		0.000	35.58***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 6: Regression Results

Results, 1998-2014						
Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0016616***	0.000326	0.000	0.001682***	0.0003071	0.000
Share of Population 20-29 (%)	0.0259725	0.0199512	0.193	**0.0296462***	0.0139923	0.063
Share of Population 30-39 (%)	**0.0607584*** **	0.023577	0.010	0.0142119	0.0181179	0.453
Share of Population 40-49 (%)	0.0474864***	0.0170837	0.005	0.0137173	0.0129295	0.316
Share of Population 50-64 (%)	0.0546264***	0.0116922	0.000	0.0231782*	0.0114316	0.073
Unemployment Rate (%)	-0.0135857***	0.0042639	0.001	-0.0101521*	0.0048671	0.067
Number of Observations	170			170		
Wald chi2/F test	5999.02***		0.000	12.32***		0.001
Test of Equality of "population share" coefficients F test	14.51***		0.002	3.51***		0.003
Test of "population share" coefficients simultaneously equal to zero F test	24.96***		0.000	4.64**		0.026

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 6: continue

Results, 1998-2014

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.0010832***	0.0001599	0.000	0.0000467***	0.000011	0.002
Share of Population 20-29 (%)	**0.0368989*** **	0.0117585	0.002	**0.0006799**	0.0004244	0.144
Share of Population 30-39 (%)	0.0158949	0.0130626	0.224	-0.0008644	0.0009766	0.399
Share of Population 40-49 (%)	0.0176293*	0.0090481	0.051	0.0000601	0.0005949	0.233
Share of Population 50-64 (%)	0.0305322***	0.0081296	0.000	0.0001756	0.0005029	0.735
Unemployment Rate (%)	-0.0047065	0.0048369	0.331	-0.0005606*	0.0002896	0.085
Number of Observations	170			170		
Wald chi2/F test	372.03***		0.000	15.58***		0.000
Test of Equality of "population share" coefficients F test	12.5***		0.006	6.64***		0.002
Test of "population share" coefficients simultaneously equal to zero F test	58.8***		0.000	6.73***		0.008

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

IV. Results and Discussion

Table 2 shows the results of our estimations from 1988-2014. Panel A reports the results of the fixed-effects Poisson regressions with cluster (province) robust standard errors. Panel B reports the results when we used a standard fixed-effects OLS regression method with cluster (province) robust standard errors instead of a count data method. In the latter case, the dependent variable is the natural logarithm of the total establishments for checking the robustness of the results from the Poisson regressions[5]. Panel C reports the results of the fixed-effects Negative Binominal regressions with cluster (province). Panel D reports the results when we used a standard fixed-effects OLS regression method with cluster (province) robust standard errors. The dependent variable in this case is the establishments per capita. Regression results show that the coefficient estimates from the four estimation methods are not significantly different from each other (i.e., the dominant age group of entrepreneurship is 40-49). In both specifications of Panel A and Panel B, the coefficients of populations share by age cohort can be arranged along an inverted U-shape with increasing age. The coefficient reaches its maximum of 0.125 for the population share of the cohort aged 40-49, meaning that a 1-percentage-point increase in the population share aged 40 to 49 increases the number of establishments by 12.5% (or 0.1-percentage-point increase in the population share aged 40 to 49 increases the number of establishments by 1.25%, which is more

[5] See Appendix 1.3.

realistic about the data on population share) --- an economically significant amount. The hypotheses that the equality of the population share coefficients and the population share coefficient are jointly zero are both rejected at the 1% level of significance, indicating that the effects of age shares on entrepreneurial activities are significant and different across the age groups. Not surprisingly, the results also show that our control variables --- both the population size and unemployment rate are significant determinants of the total establishments in each province. In particular, an increase of 10,000 in total population increases total establishments by 0.26% and a 1-percentage-point increase in the unemployment decreases the total establishments by 6.7%, indicating strong regional effects on business activities. In order to be more robust, we also used fixed-effects negative binomial regression to control overdispersion and standard fixed-effects OLS regression with dependent variable establishments per capita to control heteroscedasticity and the effect of time trend (i.e., both establishments and population data have the same trend over time), but the results from Panel C and Panel D are not significantly different. Moreover, this finding is consistent with what we could see in Figure 3: during the overall time period, the cohorts aged 40-49 play a dominant role in entrepreneurial activities.

To find out whether the relationships between provincial age structure and business establishment activity is stable over time (i.e., during demographic change), we split our sample period and ran the regressions separately. Although this procedure will result in the loss of with-in variation of our data, it could give us a first indication of

the stability of our findings associated with demographic change. Table 3 demonstrates the results of regressions for 1988 to 2000 and Table 4 contains the results for 2001 to 2014. The estimates reveal that the general inverse U-shaped relationship between age structure and total establishments continue to hold in the period 1988-2000, with its maximum at ages 40 to 49, which is consistent with our findings of the overall time period. However, a peak occurred for ages 20 to 29 in the period 2001-2014. This finding provides us with strong evidence of the "age-specific peer effects". As is suggested in Bronte, Falck & Heblich (2009), we cannot simply multiply the size of an age cohort with a given age-specific likelihood of becoming an entrepreneur to arrive at the expected number of establishments in an age cohort because the age-specific likelihood of starting a business changes with the size of the age cohort. With the development of Internet Technology in recent years, entrepreneurs could benefit more from knowledge spillovers, internet-based external finance and their social networks. Thus, younger generations could easily adopt the Internet Technology to start their own businesses.

In addition, Table 5 and Table 6 present the regression results for 1988-1997 and 1998-2014 respectively. In accordance with our previous findings, the general inversed U-shape continues to hold and the dominant age groups that could promote

entrepreneurial activities are 40-49 for 1988-1997 and 20-29 for 1998-2014[6], indicating the existence of "age-specific peer effects" as well.

Furthermore, as shown by the F-tests of the hypothesis that the equality of the population share coefficients and the population share coefficient are jointly zero are both rejected at the 1% and 5% levels of significance in both periods indicating that different age-structures do have effects on entrepreneurial activities.

As a check of robustness, we also considered the hysteresis effect since the dependent variable we used, establishment, is a stock variable. Also any adjustment in entrepreneurial activity would be slower to show up in the number of total establishments. To overcome this problem, we include the one-year-lagged establishment as the independent variable[7]. Results show that the dominant age cohorts that could promote entrepreneurial activities still remain the same as our previous findings, and the general inversed U-shaped relationship between age structure and entrepreneurship continues to hold in the time period 1988-2014 and 1988-2000.

V. Conclusions

Due to a sustained decline in mortality and fertility rates in most Western countries, there has been a steady shift towards older populations worldwide. Ageing societies

[6] From Table 6, Panel A reports cohort aged 30-39 is the dominant age group whereas Panel B, Panel C and Panel D report cohort aged 20-29 is the dominant age group.
[7] See Appendix 2.1, Additional Regression Results.

are usually considered problem cases due to rising health care costs, falling productivity and a stagnation of economic dynamics. Literature has identified a close link between age structure and startup activities using micro---level data in many countries including US, New Zealand and Australia. Literature has also identified that entrepreneurship is a regional phenomenon which requires the policy attention of all regional governments.

Based on the research conducted by Bronte, Falck & Heblich (2009) for Germany, a count data model and a longitudinal data set for Canada at the aggregate level were used to analyze the the impact of provincial age structure on the aggregate entrepreneurial activities in each of the ten Canadian provinces from 1988 to 2014. We found the inverse U-shaped relationship between age cohort and total establishments in ten Canadian provinces, which is accord with Bronte, Falck & Heblich (2009). Moreover, this finding is consistent with previous empirical studies based on cross-sectional individual data. Since our sample period from 1988 to 2014 covers both the early years and later years of the baby boomer cohorts, we were able to examine the changes of this inverse U-shaped relationship between provincial age structure and entrepreneurial activity by splitting our sample into two subsamples: from 1988 to 2000 and from 2001 to 2014. We found that while the general pattern of an inverse U-shaped relationship holds in the period 1988 to 2000, the peak, however, moved from the 40-49 age group from 1988 to 2000 to the 20-29 age cohort from 2001-2014, reflecting the existence of age-specific peer effects during

the sample period. Not surprisingly, when splitting our sample in the year of 1997, the results are consistent with the splitting in the year of 2000, which also indicated the age-specific peer effects.

Contrary to popular perception but consistent with the literature, our results suggest that entrepreneurship is not exclusive to the young and hip. Entrepreneurs of all ages start businesses and create economic opportunities for themselves and others.

However, what could not be overlooked is the fact that the entrepreneurial growth peak has been moved from the older groups aged about 40-49 before 21st century to the younger groups aged around 20-29 in recently years under the influence of the peer effects. It is obvious that the younger generations are generally taking the lead, so considering that the impact of age structure on the entrepreneurship is observable and measurable, proper and useful policies that could facilitate more start-ups for the younger groups are essential. For example, policy-oriented start-up support plans by local governments will have a great effect on people who are planning on a new business.

In another aspect, although the seniors are no more the major groups to start a business, that does not mean their value is limited nowadays. The seniors, on the other hand, possess more experience, better social networking, and greater financial resources, which play a great role in starting a new business. Therefore, policymakers can support these “third age” or “encore” entrepreneurs by perfecting

policies that lower barriers to entrepreneurial entry, maintain flexible labour markets, and establish government-oriented seed funds. At the same time, the phenomenon of ageing society is inevitable. Since every young man will get their wrinkles on the faces, some measures should be taken to make a good transfer of some knowledge and skills from the seniors to the juniors such as more intergenerational interactions. And thus, the older could get access to the development of the society and the economy. In a recent article published in The Globe and Mail (Sept 24, 2014): The upside to Canada's ageing population, the author suggests several reasons to be optimistic about the impact of an ageing society on the Canadian economy including better health, longer life expectancy, higher education, reduced financial burden on young people, better environment and increased quality of life. From this perspective, deeper research remains to be done to understand more details of the influence of the peer effects in order that proper policies can be made to fully perform their roles in an ageing society. However, as mentioned before this study is subject to a number of limitations: First, the dependent variable we used is the count of total establishments, seeing that we cannot get the whole 27-year data on establishment size, it is hard to make sure that the average establishment size is stable over time. In this case, we assume implicitly that more smaller firms are somehow better than fewer larger firms even though it may not be the case. Second, it appears to be a break in the series of establishment counts in 1997, where the number of establishments increased in all provinces by about 50%, based on Figure 2. The reason could be the change of

statistical definitions just as the Statistics Canada states, “some establishments do not employ any individuals directly”. Finally, there is a possibility that establishments, population and unemployment rate are simultaneously determined, since prospering regions are increasing in population, attracting more people and offering more entrepreneurial opportunities, hence lower the unemployment rate at the same time. Ideally, we need to find proper instrument variables to isolate the exogenous variations of population and unemployment rate. However, since the dependent variable we used is a stock variable, the mission is almost impossible. Thus, we could measure the impact of regional age structure on entrepreneurship in the single way we discussed above, without making an improper implicit assumption that there exist causal relationships between those variables.

VI. References

Angrist, J. D., & Krueger, A. B. (2001). Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, *15*(4), 69-85.

Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, *86*(3), 630-640.

Blanchflower, D. G. (2004). *Self-employment: More may not be better* (No. w10286). National Bureau of Economic Research.

Blundell, R., Griffith, R., & Van Reenen, J. (1995). Dynamic Count Data Models of Technological Innovation. *The Economic Journal*, 333-344.

Bonte, W., Falck, O., & Heblich, S. (2007). Demography and Innovative Entrepreneurship. *CESifo Working Paper,* No. 2115.

Carree, M. A., & Thurik, A. R. (2010). The Impact of Entrepreneurship on Economic Growth. In *Handbook of Entrepreneurship Research* (pp. 557-594). Springer US.

Chen, Q. (2013). Count Data Models. *Advanced Econometrics and Stata Application*. (pp.211-222). China, Beijing: Higher Education Press.

Evans, D. S., & Leighton, L. S. (1989). Some Empirical Aspects of Entrepreneurship. *The American Economic Review*, *79*(3), 519-535.

Freytag, A., & Thurik, R. (2007). Entrepreneurship and its Determinants in a Cross-country Setting. *Journal of Evolutionary Economics*, *17*(2), 117-131.

Gartner, W. B. (1985). A Conceptual Framework for Describing the Phenomenon of New Venture Creation. *Academy of Management Review*, *10*(4), 696-706.

Ghani, E., Kerr, W. R., & O'connell, S. (2011). Spatial Determinants of Entrepreneurship in India. *Harvard Business School, Working Paper,* 12-027.

Grilo, I., & Thurik, R. (2008). Determinants of Entrepreneurial Engagement Levels in Europe and the US. *Industrial and Corporate Change*, *17*(6), 1113-1145.

Guiso, L., Sapienza, P., & Zingales, L. (2004). The Role of Social Capital in Financial Development. *The American Economic Review*, *94*(3), 526-556.

Hughes, A. (2003). Knowledge Transfer, Entrepreneurship and Economic Growth: Some Reflections and Implications for Policy in the Netherlands. ESRC Centre for Business Research, *University of Cambridge, Working Paper,* No. 273.

Iyigun, M., & Owen, A. L. (1997). Risk, Entrepreneurship and Human Capital Accumulation. *Board of Governors of the Federal Reserve System Finance and Econ. Disc. Series*, 97-37.

Kurek, S., & Rachwał, T. (2011). Development of Entrepreneurship in Ageing Populations of the European Union. *Procedia-Social and Behavioral Sciences*, *19*, 397-405.

Levesque, M., & Minniti, M. (2006). The Effect of Aging on Entrepreneurial Behavior. *Journal of Business Venturing*, *21*(2), 177-194.

Li, H., Yang, Z., Yao, X., & Zhang, J. (2009). Entrepreneurship and Growth: Evidence from China. *Discussion papers*, 22.

Manski, C. F. (2000). *Economic analysis of social interactions* (No. w7580). National Bureau of Economic Research.

Meyer, B. D., & Blanchflower, D. G. (1991). *A longitudinal analysis of young entrepreneurs in Australia and the United States*. National Bureau of Economic Research.

Parker, S. C. (2004). *The Economics of Self-employment and Entrepreneurship*. Cambridge University Press.

Schumpeter, J. A. (1951). *Essays: On Entrepreneurs, Innovations, Business cycles, and the Evolution of Capitalism*. Transaction Publishers.

Stam, E. (2007). Why Butterflies Don't Leave: Locational Behavior of Entrepreneurial Firms. *Economic Geography*, 27-50.

Stangler, D., & Spulber, D. F. (2013). The Age of the Entrepreneur: Demographics and Entrepreneurship. Retrieved from http://iiij.org/wp-content/uploads/2013/05/i4jDaneStanglerDemographicsandEntrepreneurship-1.pdf

Toth, R. (2012). Entrepreneurial Human Capital and Entrepreneurial Dynamics: Evidence from a Natural Experiment. Retrieved from http://www.iza.org/conference_files/EntreRes2012/toth_r7741.pdf

Wang, C., Chen, X., & Zhu, X. (2014). Age Structure and Entrepreneurship: Evidence from China Provincial Level Data. *Contemporary Economic Research, 7*, 013.

Weber, P., & Schaper, M. (2003, September). Understanding the Grey Entrepreneur: A Review of the Literature. In *the16th Annual Conference of Small Enterprise Association of Australia and New Zealand, University of Ballarant, Victoria, Australia*.

Wong, P. K., Ho, Y. P., & Autio, E. (2005). Entrepreneurship, Innovation and Economic growth: Evidence from GEM data. *Small Business Economics*, *24*(3), 335-350.

Zhang, T. (2008). *Elderly Entrepreneurship in an Aging US Economy: It's never too late* (Vol. 2). World Scientific.

Appendix

1.1 Poisson Regression Model

For individual *i*, denote dependent variable as Y_i, assuming the probability of $Y_i = y_i$ is determined by a Poisson distribution with coefficient λ_r:

$$P(Y_i = y_i \mid \boldsymbol{x_i}) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!} \quad (y_i = 0, 1, 2, \ldots)$$

where $\lambda_i > 0$ is the Poisson arrival rate, which means the average frequency of Y happened is determined by the explanatory variable $\boldsymbol{x_i}$ and

$E(Y_i \mid \boldsymbol{x_i}) = \mathrm{Var}(Y_i \mid \boldsymbol{x_i}) = \lambda_i$.

In order to keep $\lambda_i \geq 0$, the conditional mean function of Y_i could be written as

$E(Y_i \mid \boldsymbol{x_i}) = \lambda_i = \exp(\boldsymbol{x_i'\beta})$. Therefore, $\ln \lambda_i = \boldsymbol{x_i'\beta}$, is a logarithm linear model.

Assuming the sample is independent and identically distributed, the sample's likehood function will be $L(\boldsymbol{\beta}) = \frac{\exp(-\sum_{i=1}^{n}\lambda_i)\cdot\prod_{i=1}^{n}\lambda_i^{y_i}}{\prod_{r=1}^{n} y_r!}$,

$\ln(\boldsymbol{\beta}) = \sum_{i=1}^{n}[-\exp(\boldsymbol{x_i'\beta}) + y_i\boldsymbol{x_i'\beta} - \ln(y_i!)]$.

First order condition is $\sum_{i=1}^{n}[y_r - \exp(\boldsymbol{x_i'\beta})]\boldsymbol{x_i} = \boldsymbol{0}$. Through this first order condition, we could derive $\widehat{\boldsymbol{\beta}}_{\mathrm{MLE}}$.

1.2 Fixed-effects Model

The general individual-specific effects model could be written as:

$$y_{it} = \boldsymbol{x_{it}'\beta} + \boldsymbol{z_i'\delta} + u_i + \varepsilon_{it} \ (i = 1, \ldots, n\,; t = 1, \ldots, T)$$

where $\boldsymbol{z_i}$ denotes time invariant individual characteristics (i.e. $\boldsymbol{z_{it}} = \boldsymbol{z_i}, \forall t$); $\boldsymbol{x_{it}}$ is time-varying. In the composite error term $u_i + \varepsilon_{it}$, u_i denotes unobserved

individual heterogeneity, ε_{it} satisfies all the classical error term properties and is uncorrelated with u_i. If u_i is correlated with one of the explanatory variables, using OLS to estimate the regression equation will yield inconsistent estimators. Take the time weighting average on both sides of the general individual-specific effects model will give us: $\overline{y_i} = \overline{\boldsymbol{x}_i'}\boldsymbol{\beta} + \overline{\boldsymbol{z}_i'}\boldsymbol{\delta} + u_i + \overline{\varepsilon_i}$. Time demeaning the original model and define $\tilde{y}_{it} \equiv y_{it} - \overline{y_i}$, $\tilde{\boldsymbol{x}}_{\boldsymbol{it}} \equiv \boldsymbol{x}_{\boldsymbol{it}} - \overline{\boldsymbol{x}_{\boldsymbol{i}}}$, $\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{it}} \equiv \boldsymbol{\varepsilon}_{\boldsymbol{it}} - \overline{\boldsymbol{\varepsilon}_{\boldsymbol{i}}}$, we could get $\tilde{y}_{it} = \tilde{\boldsymbol{x}}_{\boldsymbol{it}}'\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{it}}$. If $\tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{it}}$ is not correlated with $\tilde{\boldsymbol{x}}_{\boldsymbol{it}}$, we could use OLS to get consistent fixed effects estimator $\widehat{\boldsymbol{\beta}}_{FE}$.

1.3 Poisson Regression Model vs OLS Regression Model

The Poisson Regression Model we used in this report is:

E (*Establishments* $_{it}$ | *Population* $_{it}$, *AgeStructure* $_{it}$, *UnemploymentRate* $_{it}$)

$= \lambda_{it} = \exp(\alpha_i + \beta_1$ *Population* $_{it} + \beta_2$ *AgeStructure* $_{it} + \beta_3$ *Unemployment rate* $_{it})$

In order to be comparable (i.e. a check of robustness), we could take the natural logarithm form of the above regression equation, then the regression equation becomes:

ln Establishments $_{it} = \alpha_i + \beta_1$ *Population* $_{it} + \beta_2$ *AgeStructure* $_{it} + \beta_3$ *Unemployment rate*$_{it}$.

2.1 Additional Regression Results

Table 7: Regression Results

Results, 1988-2014						
Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0001401	0.0006258	0.823	0.0012163	0.0009457	0.230
Share of Population 20-29 (%)	0.0044517	0.0293661	0.880	0.0516406	0.0310284	0.130
Share of Population 30-39 (%)	0.1073077***	0.0255788	0.000	-0.0005828	0.0361264	0.987
Share of Population 40-49 (%)	**0.1096537*** **	0.0295232	0.000	**0.1249238*** **	0.0265331	0.001
Share of Population 50-64 (%)	0.0127327***	0.0217471	0.000	0.0505744**	0.0218825	0.046
Unemployment Rate (%)	-0.0442662***	0.0114636	0.000	-0.0342493**	0.0119695	0.019
Establishments (-1)	1.13E-06***	2.62E-07	0.000	1.02E-06***	2.60E-07	0.004
Number of Observations	260			260		
Wald chi2/F test	94454.92***		0.000	173.73***		0.000
Test of Equality of "population share" coefficients F test	55.68***		0.000	143.64***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	58.97***		0.000	107.72***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 7: continue

Results, 1988-2014

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	-0.0010588**	0.0004593	0.021	1.43E-06	0.0000371	0.970
Share of Population 20-29 (%)	0.0392002**	0.015376	0.011	0.0040721**	0.0017602	0.046
Share of Population 30-39 (%)	0.0372497*	0.0207126	0.072	-0.0008118	0.0014533	0.590
Share of Population 40-49 (%)	**0.101287*** **	0.0112804	0.000	**0.0069924*** **	0.0016957	0.003
Share of Population 50-64 (%)	0.0726493***	0.0110442	0.000	0.0023611**	0.0009374	0.033
Unemployment Rate (%)	-0.0162731*	0.0090758	0.073	-0.0015941***	0.0004884	0.010
Establishments (-1)	1.95E-06***	2.95E-07	0.000	4.96E-08***	1.54E-08	0.010
Number of Observations	260			260		
Wald chi2/F test	1449.28***		0.000	76.68***		0.000
Test of Equality of "population share" coefficients F test	44.04***		0.000	78.51***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	246.81***		0.000	58.08***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 8: Regression Results

Results, 1988-2000

Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.004644***	0.0006641	0.000	0.0030905***	0.0008363	0.005
Share of Population 20-29 (%)	0.0290616	0.0216367	0.179	-0.0026928	0.0382716	0.945
Share of Population 30-39 (%)	-0.239683***	0.0333568	0.000	-0.2996532***	0.0538516	0.000
Share of Population 40-49 (%)	**0.1020326*** **	0.033416	0.002	**0.1021534*** **	0.0286544	0.006
Share of Population 50-64 (%)	0.028495	0.0249911	0.254	-0.0124878	0.0266591	0.651
Unemployment Rate (%)	-0.0132549	0.0097482	0.174	0.0002345	0.0108331	0.983
Establishments (-1)	5.09E-08	1.22E-07	0.676	3.36E-07	2.66E-07	0.237
Number of Observations	120			120		
Wald chi2/F test	800956.23***		0.000	318.43***		0.000
Test of Equality of "population share" coefficients F test	206.23***		0.000	33.22***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	407.96***		0.000	76.28***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 8: continue

Results, 1988-2000

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	-0.000222	0.0004807	0.644	0.0000744	0.0000571	0.225
Share of Population 20-29 (%)	-0.0451355*	0.0254663	0.076	0.0008466	0.0022692	0.718
Share of Population 30-39 (%)	-0.321133***	0.0430143	0.000	-0.0174686***	0.0033705	0.001
Share of Population 40-49 (%)	**0.0949903*** **	0.0252164	0.000	**0.0054394** **	0.0021833	0.034
Share of Population 50-64 (%)	-0.0421223**	0.0188881	0.026	-0.0014653	0.0014932	0.352
Unemployment Rate (%)	0.0088536	0.0110687	0.424	0.0004583	0.0005908	0.458
Establishments (-1)	8.55E-07***	2.32E-07	0.000	1.40E-08	1.66E-08	0.421
Number of Observations	120			120		
Wald chi2/F test	1154.81***		0.000	64.59***		0.000
Test of Equality of "population share" coefficients F test	213.38***		0.000	22.88***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	404.77***		0.000	39.00***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 9: Regression Results

Results, 2001-2014						
Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	0.0001418	0.0004265	0.740	-0.000142	0.0003121	0.660
Share of Population 20-29 (%)	**0.0286495***	0.0155428	0.065	**0.0164843**	0.0144009	0.282
Share of Population 30-39 (%)	0.0284097	0.0180498	0.115	0.0070988	0.0135118	0.612
Share of Population 40-49 (%)	-0.0190885	0.0140694	0.175	-0.039635***	0.0091722	0.002
Share of Population 50-64 (%)	0.0080505	0.0120014	0.502	-0.0088645	0.010258	0.410
Unemployment Rate (%)	-0.0060252***	0.0022888	0.008	-0.0084251**	0.0027097	0.013
Establishments (-1)	9.71E-07***	3.00E-07	0.001	1.31E-06***	3.33E-07	0.003
Number of Observations	130			130		
Wald chi2/F test	6227.67***		0.000	19.32***		0.001
Test of Equality of "population share" coefficients F test	58.58***		0.000	37.10***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	56.98***		0.000	28.64***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 9: continue

Results, 2001-2014

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	-0.0002914	0.0002369	0.219	-0.0000607**	0.0000224	0.024
Share of Population 20-29 (%)	**0.0166451***	0.00938	0.076	**3.69E-06**	0.0006749	0.996
Share of Population 30-39 (%)	0.0116616	0.0081718	0.154	-0.0016305***	0.0004172	0.004
Share of Population 40-49 (%)	-0.0369983***	0.0072064	0.000	-0.0027165***	0.0003771	0.000
Share of Population 50-64 (%)	-0.0037293	0.0063947	0.560	-0.0020856***	0.0003663	0.000
Unemployment Rate (%)	-0.006321**	0.0030167	0.036	-0.0004595***	0.0001079	0.002
Establishments (-1)	1.22E-06***	2.00E-07	0.000	8.45E-08***	2.06E-08	0.003
Number of Observations	130			130		
Wald chi2/F test	386.17***		0.000	20.03***		0.000
Test of Equality of "population share" coefficients F test	105.90***		0.000	13.33***		0.001
Test of "population share" coefficients simultaneously equal to zero F test	110.43***		0.000	17.02***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 10: Regression Results

Results, 1988-1997

Method	Panel A: Dependent Variable=Establishments FE Poisson Regression with Cluster (Province)			Panel B: Dependent Variable=log(Establishments) FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.0052804***	0.0005415	0.000	0.0037074**	0.0011915	0.012
Share of Population 20-29 (%)	0.0475056**	0.0194033	0.014	0.0052881	0.0219767	0.815
Share of Population 30-39 (%)	-0.2599095***	0.0451104	0.000	-0.2535493***	0.0379345	0.000
Share of Population 40-49 (%)	**0.1250945*** **	0.0234427	0.000	**0.1028033*** **	0.0192744	0.000
Share of Population 50-64 (%)	0.0475696	0.0543011	0.381	-0.0119182	0.0302678	0.703
Unemployment Rate (%)	0.0016583	0.0023387	0.478	0.0055229	0.005224	0.318
Establishments (-1)	-4.62E-06***	1.60E-06	0.004	-2.71E-06	2.80E-06	0.357
Number of Observations	90			90		
Wald chi2/F test	24012.83***		0.000	72.27***		0.000
Test of Equality of "population share" coefficients F test	164.53***		0.000	37.43***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	188.33***		0.000	40.04***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 10: continue

Results, 1988-1997

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	0.0005259	0.0011359	0.643	0.0001618***	0.0000434	0.005
Share of Population 20-29 (%)	-0.0099984	0.0229935	0.664	-0.0003697	0.0007631	0.640
Share of Population 30-39 (%)	-0.2666982***	0.0561135	0.000	-0.0135613***	0.0019052	0.000
Share of Population 40-49 (%)	**0.1015829*** **	0.0252119	0.000	**0.0038855*** **	0.0007026	0.000
Share of Population 50-64 (%)	-0.0308054	0.0278104	0.268	-0.0012611	0.0009305	0.208
Unemployment Rate (%)	0.0100509	0.0091396	0.271	0.0003968	0.0002212	0.106
Establishments (-1)	8.44E-07	3.01E-06	0.779	-3.03E-07**	1.01E-07	0.015
Number of Observations	90			90		
Wald chi2/F test	216.81***		0.000	263.61***		0.000
Test of Equality of "population share" coefficients F test	60.71***		0.000	38.66***		0.000
Test of "population share" coefficients simultaneously equal to zero F test	97.92***		0.000	38.16***		0.000

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 11: Regression Results

Results, 1998-2014						
Method	**Panel A: Dependent Variable=Establishments**			**Panel B: Dependent Variable=log(Establishments)**		
	FE Poisson Regression with Cluster (Province)			**FE OLS Regression with Cluster (Province)**		
	Coefficients	**Robust Standard Errors**	**p-values**	**Coefficients**	**Robust Standard Errors**	**p-values**
Population Size (in 10,000)	-0.0002767	0.0005599	0.621	-0.0003699	0.0003484	0.316
Share of Population 20-29 (%)	0.02351	0.0144642	0.104	0.018225	0.0149937	0.255
Share of Population 30-39 (%)	**0.0430171****	0.0182688	0.019	**0.0183556**	0.0150258	0.253
Share of Population 40-49 (%)	0.0015088	0.0158257	0.924	-0.0147243	0.0097975	0.167
Share of Population 50-64 (%)	0.0341723***	0.0120461	0.005	0.0142892	0.0102197	0.196
Unemployment Rate (%)	-0.005406*	0.0030376	0.075	-0.0067643*	0.0030856	0.056
Establishments (-1)	1.22E-06***	3.29E-07	0.000	1.46E-06***	3.30E-07	0.002
Number of Observations	160			160		
Wald chi2/F test	54269.75***		0.000	25.62***		0.001
Test of Equality of "population share" coefficients F test	55.71***		0.000	17.10***		0.001
Test of "population share" coefficients simultaneously equal to zero F test	57.23***		0.000	12.83***		0.001

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

Table 11: continue

Results, 1998-2014

Method	Panel C: Dependent Variable=Establishments			Panel D: Dependent Variable=Establishments per capita		
	FE NB Regression with Cluster (Province)			FE OLS Regression with Cluster (Province)		
	Coefficients	Robust Standard Errors	p-values	Coefficients	Robust Standard Errors	p-values
Population Size (in 10,000)	-0.0005392*	0.0002639	0.041	-0.0000812**	0.0000267	0.014
Share of Population 20-29 (%)	0.0181366**	0.0087606	0.038	**0.0001517**	0.0006151	0.811
Share of Population 30-39 (%)	**0.0225393****	0.008998	0.012	-0.0006995	0.0006285	0.295
Share of Population 40-49 (%)	-0.0140715**	0.0067682	0.038	-0.0010856***	0.000336	0.010
Share of Population 50-64 (%)	0.0180236***	0.00577	0.002	-0.0004578	0.0003163	0.182
Unemployment Rate (%)	-0.0049605	0.0034894	0.155	-0.0003547*	0.000164	0.059
Establishments (-1)	1.47E-06***	1.99E-07	0.000	9.31E-08***	2.46E-08	0.004
Number of Observations	160			160		
Wald chi2/F test	506.06***		0.000	23.23***		0.000
Test of Equality of "population share" coefficients F test	70.49***		0.000	3.32*		0.071
Test of "population share" coefficients simultaneously equal to zero F test	96.64***		0.000	2.98*		0.080

Notes: *Statistically significant at the 10% level; **statistically significant at the 5% level; *** statistically significant at the 1% level.

CURRICULUM VITAE

Candidate's Full Name: Liang Ma

University Attended: Bachelor of Economics,
Shandong University of Finance and Economics, 2015.7.

Publications: N.A.

Conference Presentations: N.A.