User Entrepreneurship and Firm Employment Growth

Zheng Tian¹²; Luyi Han¹²; Timothy R. Wojan³; Anil Rupasingha⁴; Stephan J. Goetz¹²

¹Pennsylvania State University, ²Northeast Regional Center for Rural Development, ³Oakridge Institute for Science and Education Fellow at the National Center for Science and Engineering Statistics, ⁴USDA Economic Research Service

Abstract

This study examines the employment growth of firms founded by user entrepreneurs in comparison with other types of entrepreneurs. In contrast to conventional entrepreneurship that individuals or businesses develop new products to fill latent demands in the market, user entrepreneurs develop products to satisfy their own needs and enter markets only after the utility has been demonstrated. Using the establishment-level microdata from the 2014 Rural Establishment Innovation Survey and the Longitudinal Business Database, we estimate the growth trajectories of firms with various types of entrepreneurship origins using random coefficient and quantile regression models. We find that firms founded by user entrepreneurs have higher employment growth rates in the first five years, but slower growth rates in subsequent years. Our study confirms the contribution of user entrepreneurship to job creation and calls for explicit acknowledgement and identification that to date has been lacking in entrepreneurship evaluation and promotion programs.

Introduction

Research questions
What are the employment growth trajectories of firms founded by user entrepreneurs? Are they different than those founded by conventional or other types of entrepreneurs?

Who are user entrepreneurs?
"The commercialization of a new product and/or service by an individual or group of individuals who are also users of that product and/or service." (Shah and Tripsas 2007, p.124)

Our contribution to the literature
- This is the first paper using econometric models to evaluate the growth performance of firms founded by user entrepreneurs.
- Previous studies on user entrepreneurship were mostly qualitative, except Shah, Smith, and Reedy (2012) using the Kauffman Firm Survey (KFS).

Data

The 2014 Rural Establishment Innovation Survey (REIS)
The REIS is the first nationally representative sample of self-reported innovation and user entrepreneurship in the tradable rural nonfarm sector (Wojan and Parker 2017).

The 1990-2019 Longitudinal Business Database (LBD)
- The LBD is a confidential database with microdata for private, nonfarm business establishments, accessed via the Federal Statistical Research Data Center (FSRDC).
- We linked establishments in the 2014 REIS with their records in the 1990-2019 LBD, matched by EIN, 3-digit NAICS, and FIPS.

Sample selection
- We confined the samples to single-unit establishments to avoid confounding factors, e.g., location, management, and the distribution effect of branches, and to limit our focus on small or medium-sized firms.
- We choose firms established after 1990 and before 2014 to observe their full employment history given data availability.

Potential problem in sample selection
- We cannot observe the survivability of establishments with different types of entrepreneurs who had exited the business before 2014 when the REIS was conducted.
- We need to control heterogeneity in initial conditions of founding and growth potentials over a long period of 1990-2014.

Methods

We divide the samples into three cohorts to mitigate the concern over heterogeneity of initial conditions and growth potentials.

We use two modeling methods – random coefficient growth models and quantile regression models – to account for firm heterogeneity and the right-tailed growth distribution.
- Random coefficient growth model (multilevel model)
  - Level 1 (establishment): \( \ln y_{it} = \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \beta_{3i}t^3 + \varepsilon_{it} \)
  - Level 2 (establishment): \( \beta_{0i} = r_0 + \gamma_i \), \( \beta_{1i} = r_1 + \gamma_i \), \( \beta_{2i} = r_2 + \gamma_i \), \( \beta_{3i} = r_3 + \gamma_i \), for \( i = 0, 1, 2, 3 \)
  - Random effects: \( \gamma_i \)
  - AR(1) within-group error: \( \varepsilon_{it} = \phi \varepsilon_{it-1} + \varepsilon_{it} \), \( \phi = N(0,0.2) \)
  - Quantile regression model: \( Q_i(\ln y_{it}) (X_i) = X_i \beta_i(r) \), for \( r \in (0,1) \), \( t = 1, ..., T \)  
  - \( y_{it} \) is the level of an establishment's employment; \( X_i \) includes entrepreneur types, initial employment size, industry (3-digit NAICS) and state fixed effects; \( \ln y_{it} = \ln y_{it} - \ln y_{i-4} \), the five-year growth rate of establishment's employment.

Data

The 2014 Rural Establishment Innovation Survey (REIS)
The REIS is the first nationally representative sample of self-reported innovation and user entrepreneurship in the tradable rural nonfarm sector (Wojan and Parker 2017).

The 1990-2019 Longitudinal Business Database (LBD)
- The LBD is a confidential database with microdata for private, nonfarm business establishments, accessed via the Federal Statistical Research Data Center (FSRDC).
- We linked establishments in the 2014 REIS with their records in the 1990-2019 LBD, matched by EIN, 3-digit NAICS, and FIPS.

Sample selection
- We confined the samples to single-unit establishments to avoid confounding factors, e.g., location, management, and the distribution effect of branches, and to limit our focus on small or medium-sized firms.
- We choose firms established after 1990 and before 2014 to observe their full employment history given data availability.

Potential problem in sample selection
- We cannot observe the survivability of establishments with different types of entrepreneurs who had exited the business before 2014 when the REIS was conducted.
- We need to control heterogeneity in initial conditions of founding and growth potentials over a long period of 1990-2014.

Methods

We divide the samples into three cohorts to mitigate the concern over heterogeneity of initial conditions and growth potentials.

We use two modeling methods – random coefficient growth models and quantile regression models – to account for firm heterogeneity and the right-tailed growth distribution.
- Random coefficient growth model (multilevel model)
  - Level 1 (establishment): \( \ln y_{it} = \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \beta_{3i}t^3 + \varepsilon_{it} \)
  - Level 2 (establishment): \( \beta_{0i} = r_0 + \gamma_i \), \( \beta_{1i} = r_1 + \gamma_i \), \( \beta_{2i} = r_2 + \gamma_i \), \( \beta_{3i} = r_3 + \gamma_i \), for \( i = 0, 1, 2, 3 \)
  - Random effects: \( \gamma_i \)
  - AR(1) within-group error: \( \varepsilon_{it} = \phi \varepsilon_{it-1} + \varepsilon_{it} \), \( \phi = N(0,0.2) \)
  - Quantile regression model: \( Q_i(\ln y_{it}) (X_i) = X_i \beta_i(r) \), for \( r \in (0,1) \), \( t = 1, ..., T \)  
  - \( y_{it} \) is the level of an establishment's employment; \( X_i \) includes entrepreneur types, initial employment size, industry (3-digit NAICS) and state fixed effects; \( \ln y_{it} = \ln y_{it} - \ln y_{i-4} \), the five-year growth rate of establishment's employment.

Conclusions

Firms with user entrepreneurship tend to have faster employment growth in the first five years than those with other types of entrepreneurship, especially at the upper quartile of firm employment growth distribution.

The REIS/KFS question is included in the provisional 2024 ABS survey that will provide much larger samples and help differentiate entrepreneurial firms introducing new products from businesses founded to supply standard products.

Disclaimers and Acknowledgements

The findings and conclusions in this presentation are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. This presentation provides results of exploratory research sponsored in part by the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF). This information is being shared to inform interested parties of ongoing activities and to encourage further discussion. Any views expressed are those of the authors and not necessarily those of NCSES, NSF, or the U.S. Census Bureau. The Census Bureau’s Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2083, Census Bureau Disclosure Approval Number CBDR8-FY13-0124. This work was supported in part by the Department of Agriculture (USDA), National Institute of Food and Agriculture (NIFA) under projects #2021-51150-34735, by the USDA Rural Development and Energy and Safety grant under project USDA-2018-67021-27684, and by the Pennsylvania State University Multistate/Regional Research-Appropriations under project #NE12249.

References