ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Fossen, Frank M.

Working Paper

Entrepreneurship over the Business Cycle in the United States: A Decomposition

IZA Discussion Papers, No. 12499

Provided in Cooperation with:
IZA Network @ LISER, Luxembourg Institute of Socio-Economic Research (LISER)

Suggested Citation: Fossen, Frank M. (2019) : Entrepreneurship over the Business Cycle in the United States: A Decomposition, IZA Discussion Papers, No. 12499, Institute of Labor Economics (IZA), Bonn

This Version is available at:
https://hdl.handle.net/10419/207325

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12499

Entrepreneurship over the Business Cycle in the United States: A Decomposition

Frank M. Fossen

JULY 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12499

Entrepreneurship over the Business Cycle in the United States: A Decomposition

Frank M. Fossen
University of Nevada, Reno and IZA

JULY 2019

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Entrepreneurship over the Business Cycle in the United States: A Decomposition*

Entry rates into self-employment increase during recessions and decrease during economic upswings. I show that this is mostly explained by the higher unemployment rate during a recession, together with the fact that at all times, unemployed persons have a relatively high propensity to become entrepreneurs out of necessity because they do not find paid employment. I use econometric decomposition techniques to quantify these effects based on the monthly matched US Current Population Survey before, during and after the Great Recession. I also document that this counter-cyclical pattern of entrepreneurial entry strongly applies to unincorporated entrepreneurship, but only weakly to incorporated entrepreneurship. This highlights the association of unincorporated and incorporated entrepreneurship with necessity and opportunity entrepreneurship, respectively. The results are useful for policy-makers and practitioners to understand, forecast and act on the different types of entrepreneurial activities that are to be expected over the business cycle.

JEL Classification: L26, J22, J23, M13

Keywords: entrepreneurship, business cycle, Great Recession, unemployment, opportunity, necessity, decomposition

Corresponding author:
Frank M. Fossen
Department of Economics
University of Nevada, Reno
1664 N. Virginia Street
Reno, Nevada 89557-0030
USA

E-mail: ffossen@unr.edu

* I thank Todd Sorensen for his valuable comments.

1 Introduction

Entry rates into self-employment are higher in recessions than in boom periods. In particular, the entry rate increased during the Great Recession (Fairlie, 2013). How much of this counter-cyclicality is explained by the higher unemployment rate in recessions and a generally higher individual propensity to start up out of unemployment? Alternatively, recession periods could be fundamentally different from boom periods in the sense that individual characteristics such as unemployment status have very different effects on entry into entrepreneurship at different stages of the business cycle. Policymakers and practitioners need insights to understand, forecast and act on the dynamics of entrepreneurial entry over the cycle.

In this context it is important to take into account the heterogeneity of entrepreneurs. For the purpose of this paper, the general term entrepreneur is used for self-employed individuals but is further differentiated. Although self-employment is commonly used to operationalize entrepreneurship in empirical research, this definition is very broad, and equating dynamic patterns of self-employment to innovative entrepreneurship is misleading. Therefore, to capture different entrepreneurial motivations and aspirations, such as necessity versus opportunity motives, we follow Levine and Rubinstein (2017) and distinguish between entrepreneurs starting unincorporated and incorporated businesses. Entrepreneurs with incorporated businesses tend to have stronger entrepreneurial abilities and to be more growth-oriented than entrepreneurs with unincorporated businesses (see also Shane, 2014; Herranz et al., 2017).

Unemployment pushes individuals into entrepreneurship because they experience difficulties in finding paid employment. Those who become entrepreneurs due to a lack of alternatives can be referred to as necessity entrepreneurs (Fairlie and Fossen, 2018). I will assess the relationship between necessity entrepreneurship and unincorporated entrepreneurship

by testing whether entrepreneurs coming out of unemployment are more likely to start unincorporated than incorporated businesses.

This paper contributes to the literatures by quantifying how much of the counter-cyclicality of entry into entrepreneurship can be explained by changes in the unemployment rate. The analysis is based on representative individual-level rotating panel data for the United States, the monthly matched Current Population Survey (CPS). I start from the observation that the monthly entry rate into entrepreneurship was higher during the Great Recession than before or after. I estimate that individual unemployment status increases the monthly probability of becoming an entrepreneur by about 0.9 percentage points and establish that this effect is almost constant over the business cycle. This is the first study that provides an econometric decomposition of the entry rate into entrepreneurship during different phases of the business cycle into an explained and an unexplained component. The results show that individual unemployment alone explains almost the entire counter-cyclicality of entry into entrepreneurship. I further document that the changes in entrepreneurship over the cycle are almost completely driven by changes in unincorporated entrepreneurship, which is much more responsive to unemployment than incorporated entrepreneurship. This indicates that necessity entrepreneurship is strongly and opportunity entrepreneurship only weakly counter-cyclical.

Empirical studies relating unemployment rates to entrepreneurship rates using aggregate data cannot separate the unemployment-push effect mentioned above from a reverse prosperity-pull effect and estimate a net effect, as noted by Parker (2018).[1] The unemployment-push effect describes that unemployed individuals are more likely to become entrepreneurs in order to escape unemployment (e.g., Audretsch and Vivarelli, 1996), which is sometimes also called a 'refugee effect' (Thurik et al., 2008). A prosperity-pull ef-

[1]The two effects are inseparable as well when a region-level unemployment rate is used as an explanatory variable in an analysis otherwise based on individual-level data (e.g., Henley, 2004).

fect occurs when an economic boom period increases entrepreneurship because of a high demand for products and services as well as potentially lower bankruptcy risk and, therefore, higher availability and lower costs of capital (Storey, 1991; Brünjes and Diez, 2013). Correspondingly, a recession may lead to a reverse prosperity-pull effect and decrease entrepreneurial entry. Consistent with this expectation, Bartz and Winkler (2016) provide evidence showing that the growth of young firms slowed down more than the growth of older firms in Germany during the Great Recession, and Lee and Mukoyama (2015) report that entry rates of manufacturing plants are higher in booms than in recessions in the United States. Entrepreneurship is also riskier during a recession because the option to fall back into paid employment if the business fails is more difficult.

Most of the existing empirical literature on the relationship between the dynamics of entrepreneurship and unemployment is based on time-series or country-level panel data analysis (e.g., Parker and Robson, 2004). Based on industry- and region-level panel data, Konon et al. (2018) find that entrepreneurship moves mostly with the unemployment rate in Germany, except in innovative industries with a small minimum efficient establishment size. Congregado et al. (2012) conclude from their time-series analysis that employer self-employment rates evolve pro-cyclically whereas own-account self-employment rates evolve counter-cyclically in Spain. They do not find significant associations in the United States but call for alternative measures of entrepreneurship, which we respond to by distinguishing between unincorporated and incorporated entrepreneurship.

Time-series studies by Faria et al. (2010) and Parker et al. (2012) suggest that entrepreneurship responds rapidly and substantially to cycles in unemployment. These results also suggest that entrepreneurship may in turn affect unemployment by creating (or destroying) jobs, although the effect of entrepreneurship on unemployment is estimated to be weaker and occur with a time lag (see also Fritsch and Noseleit, 2013). In this pa-

per, I focus on the contemporaneous effect of individual unemployment on the probability of becoming an entrepreneur, not the potential delayed effect of entrepreneurship on unemployment rates (see, e.g., Thurik et al., 2008). Koellinger and Thurik (2012) estimate a Vector Auto Regression (VAR) model using a panel of annual country-level data from 22 OECD countries over the period 1972-2007. The results suggest that entrepreneurship forecasts unemployment downswings one year in advance. However, Parker et al. (2012) show that estimation results from VAR models in this context are not robust to structural breaks in the data. While they find similar results from a VAR estimation as Koellinger and Thurik (2012) using quarterly aggregate data from the UK when they do not account for structural breaks, the influence of entrepreneurship on unemployment disappears when they re-estimate the VAR for a sub-period not confounded by structural breaks. I explicitly allow for the possibility of structural breaks by estimating separate coefficients for different phases of the business cycle based on individual-level panel data.

Studies using microdata are in a better position than those using aggregate data to identify the unemployment-push effect separately from the reverse prosperity-pull effect because the former effect works through an individual's unemployment status, whereas the latter effect works at the aggregate level and affects both unemployed and employed persons considering to become entrepreneurs. Furthermore, most studies using aggregate data can only consider the net entry rate, which is the difference between the entry and exit rates, whereas I can specifically estimate effects on the individual probability of entry into entrepreneurship. Studies based on individual-level data can also control for educational attainment and therefore separate effects of unemployment from effects of human capital. Papers using microdata predominantly find a positive relationship between individual unemployment and entrepreneurship, which is consistent with the unemployment-push hypothesis (Ritsilä and Tervo, 2002; Berglann et al., 2011; Åstebro et al., 2011; Biehl et al., 2014; Fritsch et al.,

2015). Studies analyzing changing patterns of transitions into entrepreneurship within business cycles using individual-level panel data are very scarce. Using the CPS, Fairlie (2013) finds that higher unemployment rates push individuals into entrepreneurship, especially out of non-employment, which is a sign of necessity entrepreneurship. He does not distinguish between unincorporated and incorporated entrepreneurship. None of the existing empirical studies decompose the change in the entry rate into entrepreneurship over the cycle into explained and unexplained components and quantify how much individual unemployment contributes to explaining the counter-cyclicality of entrepreneurial entry.

Section 2 explains the econometric method I employ, the non-linear Oaxaca-Blinder decomposition of the entrepreneurial entry rate. Section 3 describes the rotating monthly panel data I use. Section 4 presents the empirical results and extensive robustness checks, and Section 5 concludes the analysis.

2 Methodological Approach: Nonlinear Decomposition

First, I estimate logit models of the probability of becoming an entrepreneur separately for different periods before, during and after the Great Recession (GR). The binary outcome variable $entry_{(i,t+1)}$ equals 1 if individual i enters entrepreneurship between months t and $t+1$, and 0 otherwise. The latent index function of the logit model is written as

$$entry^*_{(i,t+1)} = X_{it}\beta + \epsilon_{it}, \tag{1}$$

where $entry^*$ is the propensity to enter into entrepreneurship, X is a vector of explanatory variables including a dummy variable indicating individual unemployment status, β is a coefficient vector including a constant, and ϵ is the error term. Second, I decompose the

change in the mean entry probability between periods into a part explained by changes in observed individual variables, including unemployment status, and an unexplained part reflected in changes in the coefficients and the intercept. The detailed decomposition method allows assessing the contribution of each variable of interest separately.

All the explanatory variables are observed in the month before a potential entry into entrepreneurship occurs. Among the variables in X, the main interest is in the individual unemployment status. To identify the effect *ceteris paribus*, I control for individual determinants of entrepreneurship known from the literature (e.g., Parker, 2018). It is particularly important to control for educational attainment because of its negative correlation with unemployment. I include an individual's highest educational degree obtained, age (linear and squared), gender, race, marital status, number of children, region of residence, and a dummy variable indicating whether the respondent lives in a metropolitan area. While I am able to control for the standard individual variables used in the literature on entrepreneurial choice, I might still miss relevant variables, which would increase the unexplained part in the decomposition analysis.

I implement an adaption of the decomposition approach originally suggested by Oaxaca (1973) and Blinder (1973); Fortin et al. (2011) provide an overview. Since the outcome variable is binary and I estimate logit models, I apply the weighting method for nonlinear models as described by Yun (2004), which allows for a detailed decomposition by single variables as well as coefficients.[2] In my context, the index problem discussed in the econometric decomposition literature pertains to whether the coefficients estimated for the GR period or for the comparison period should be used to assess the contribution of the variables to the change in the entrepreneurial entry rate. I follow Neumark (1988) and Oaxaca and Ransom (1994) and use the coefficients from an estimation of the logit model of entry

[2]Fairlie (2005) suggests an alternative decomposition method for nonlinear models.

based on the pooled sample including both periods. I also include a dummy variable indicating the GR period in the pooled model, as generally recommended by Jann (2008) in order to avoid a potential spillover from the unexplained part of the differential into the explained component. Furthermore, I normalize categorical variables, i.e, the educational degree, race, and regional dummy variables. As a result, effects are expressed as deviations from the overall mean, and the detailed decomposition results do not depend on the choice of an otherwise arbitrarily omitted base category (Yun, 2005). I describe this variant of a nonlinear decomposition formally in the Appendix (see also Caliendo et al., 2014).

3 Data

3.1 Representative Panel Data

For the empirical analysis, I use the monthly waves of the Current Population Survey (CPS) from January 2004 to December 2014, i.e., before, during and after the GR. The CPS is a representative survey of households in the United States provided by the Census Bureau. The U.S. Bureau of Labor Statistics relies on the CPS to estimate the widely reported national unemployment rate. The CPS follows a rotating survey design. Households are interviewed in four consecutive months, then pause for eight months, and then are surveyed again in four more consecutive months. I use the IPUMS-CPS (Flood et al., 2017), which merges these consecutive observations at the individual level to construct rotating panel data. The first three months of each four-month survey spell can be linked to the subsequent month, so 75% of all observations can be connected to the following month. Thus, for each individual, I include a maximum of six monthly observations with information on subsequent labor market transitions in our estimation sample.

The panel data structure of the matched CPS allows me to observe entries into entrepreneurship from one month to the next based on questions on the current employment

status in two consecutive months. Respondents are asked: "Last week, were you employed by government, by a private company, a nonprofit organization, or were you self-employed?" Those who respond that they were self-employed are then asked if their business is incorporated or not. In the estimation sample, I include individuals between the ages of 21 and 64 and exclude unpaid family members, those unable to work, and retirees.

For the decomposition analysis, I split the sample into three periods: before, during, and after the peak of the GR. The peak of the GR is defined as September 2008, when Lehman Brothers filed for bankruptcy, until one year later, August 2009. Figure 1 shows that this was the period of the sharp increase in unemployment in the United States. For comparison, I define the periods before and after the GR with a length of one year each as well. In the main analysis, the period before the GR is April 2007 to March 2008. This is as close as possible to the GR period, but before the increase in unemployment starts. Because the recovery after the GR was slow, I define May 2013 to April 2014 as the period after the GR. In Section 4.3, I systematically move the time periods before and after the GR month by month to assess the sensitivity of the results with respect to the definition of these periods and find that the main results are robust.

< Insert Figure 1 about here >

3.2 The Entry Rate Over The Business Cycle

Table 1 shows descriptive statistics for the individuals in the samples before, during and after the GR, as defined in the previous section for the main analysis. Before the GR, 3.7% of the individuals were unemployed. During the GR, the unemployment rate jumped up to 6.6%, before it slowly decreased to 5.5% after the GR in 2013/14.[3]

[3]Note that the term "unemployed" means that somebody is looking for paid work, and this term is different from "not in the labor force". For example, a married spouse raising children at home and currently not participating in the labor market is included in our sample, because he or she could decide to become an entrepreneur, but is not coded as unemployed.

< Insert Table 1 about here >

The share of individuals who entered into entrepreneurship between two subsequent months before the GR was 0.54%. During the GR, the monthly entry rate increased to 0.61%, then it decreased to 0.57% again. The monthly entry rate into unincorporated entrepreneurship exhibits a similar pattern, going up from 0.43% before to 0.48% during and then down again to 0.44% after the GR. In contrast, the entry rate into incorporated entrepreneurship increased slightly over the entire period.

Figure 1 shows the month-to-month entry rate into entrepreneurship in addition to the unemployment rate. Due to the low numbers of entries, these monthly averages are noisy, which is my primary motivation for pooling the monthly data over a year in the regressions. The dashed line depicts a polynomial fit of the fourth degree for the monthly entry rate. It clearly moves with the unemployment rate.

The aim of the following econometric analysis is to measure how much of the increase in the entry rate into entrepreneurship during the GR is explained by the higher unemployment rate during this period, how much can be explained by other observable factors that also changed during the GR, and how much remains unexplained. I further distinguish between unincorporated and incorporated entrepreneurship.

4 Empirical Results

4.1 Probability Model of Entry into Entrepreneurship

I first report the results of the logit estimations of the probability of entry into entrepreneurship before proceeding with the decomposition. Table 2 shows the average marginal effects of the variables on the month-to-month probability of becoming an entrepreneur for three separate logit estimations before, during and after the GR.

< Insert Table 2 about here >

In the context of this paper, the most important result is that individual current unemployment status has a strong positive effect on the probability of becoming an entrepreneur in the next month, and that this effect is almost constant over the business cycle. Before, during and after the GR, the probability of an unemployed person to become an entrepreneur was about 0.9 percentage points higher than for other persons, keeping the education level and the other controls constant. The effect is statistically significant at the 1%-level and economically very strong, as the probability of becoming an entrepreneur increases by more than 150% for the unemployed relative to the average monthly entry probabilities indicated at the bottom of the table. The effect of individual unemployment on entrepreneurial entry is stronger than that of any other variable. These results support the "unemployment-push" hypothesis.

Next, I estimate the monthly probabilities of becoming an unincorporated or an incorporated entrepreneur separately (based on the same sample of those who are not currently entrepreneurs). The results appear in Table 3. Although unemployment status has a positive and significant effect on becoming either type of entrepreneur in all periods, the effect size is five to six times larger for entry into unincorporated entrepreneurship than for entry into incorporated entrepreneurship. This confirms that running an unincorporated business is often an indicator of necessity entrepreneurship. In contrast, incorporated entrepreneurship is related to opportunity entrepreneurship, i.e., individuals typically leave their paid employment in order to become entrepreneurs directly. Another insight from this table is that the estimated effects of unemployment status on the probabilities of becoming either type of entrepreneur do not change much over the business cycle, similar to what we saw for entry into entrepreneurship in general.

< Insert Table 3 about here >

4.2 Decomposition Results

In this section I discuss the results from this paper's core analysis, the decomposition of the estimated logit models of the probability of entry into entrepreneurship. The aim is to determine how much of the difference between the entry rates during the GR versus before or after the GR can be explained by changes in the independent variables, particularly individual unemployment, and how much remains unexplained. Table 4 presents the results. The first two columns represent decompositions of the entry rate into entrepreneurship in general, including both unincorporated and incorporated entrepreneurship. The first column compares the entry rate during the GR (0.612%) to the lower entry rate before the GR (0.541%), and the second column compares it to the lower entry rate after the GR (0.568%). In both cases, the difference between the mean entry rates is significant at the 1%-level. Changes in the distributions of the independent variables (as reflected in Table 1) explain most of the difference between the entry rates during and before the GR, and a lower share, but still more than half, of the difference between the entry rates during and after the GR (see row "explained" in Table 4). In both cases, the unexplained part is not significantly different from zero. This means that changes in the coefficients and the constant over time, as reflected in Table 2, do not significantly contribute to the changes in the entry rates into entrepreneurship. The finding that the unexplained part is small also strongly suggests that I am not omitting important variables from the model, because this would increase the unexplained part.[4]

< Insert Table 4 about here >

The lower panel of the table reveals which independent variables contribute how much to the change in the entry rate. Groups of variables are considered in the cases of the education

[4]For example, think of the extreme case of not including any x-variables. Then the constants in the logit models would pick up the entire change in the entry rate, thus, 100% of the change in the entry rate would be unexplained.

and race dummies and the linear and squared age terms. The results are very clear. The change in the distribution of the dummy variable indicating current unemployment status is responsible for almost the entire explained part of the change in the entry rate, both when comparing the GR period to the periods before and after.[5] Combined with the observation that the unemployment rate was higher in the GR than before and after, the findings document that the increase in the entrepreneurial entry rate during the GR is mostly explained by the increase in unemployment. More precisely, the change in the unemployment rate explains 80% of the raw increase in the entry rate into entrepreneurship from before to the peak of the GR and 45% of the decrease thereafter.[6] The contributions of the other variables to the change in the entry rate are small.

The remaining columns of Table 4 show decompositions of the entry rates into unincorporated and incorporated entrepreneurship. The results for unincorporated entrepreneurship are similar to those for total entrepreneurship: Most of the differences in the entry rates over the business cycle can be explained by changes in the variables, and more of the increase before than of the decrease after the GR can be explained. Both before and after the GR, the unexplained part of the change in the entry rate is smaller and less significant than the explained part. Most importantly, in both cases, the unemployment dummy variable is again responsible for most of the explained part.

In contrast to unincorporated entrepreneurship, the increase in the entry rate into incorporated entrepreneurship from before the GR to the peak of it is small, and after the GR, the entry rate increases further slightly, although this further increase is not statistically significantly different from zero. Furthermore, only 40% of the increase leading up to the GR can be explained by the individual characteristics, again mostly by the unemployment status. Thus, unemployment does not play a nearly as important role for entry into

[5]To see this, compare the row labeled "Unemployed" to the row labeled "Explained".
[6]This relates the "Unemployed" row to the "Difference" row.

incorporated entrepreneurship as it does for entry into unincorporated entrepreneurship. In summary, the decomposition results document that the increase in the total entry rate into entrepreneurship during the GR is mostly due to necessity entrepreneurship out of the larger pool of unemployed individuals during the GR in the form of unincorporated entrepreneurship.

4.3 Robustness

In this section, I assess the sensitivity of the results with respect to the definition of the periods before and after the GR used for comparison with the period during the GR (09/2008, when Lehman Brothers collapsed, to 08/2009). In the main analysis above, I defined 04/2007-03/2008 as the period before the GR and 05/2013-04/2014 as the period after, as motivated in Section 3.1. Thus, the three periods all have the length of one year. In this robustness check, I conduct 71 different decompositions of the entrepreneurial entry rate exactly as described above, the only difference being that I systematically choose different comparison periods, all of length one year, over the ten years between 2004 and 2014 (Figure 2).

< Insert Figure 2 about here >

I start with a decomposition analysis using 01/2004-12/2004 as the period before the GR, then increase the start and end months by one month, using 02/2004-01/2005, and so on. Thus, I always shift the comparison time window by one month and always compare this to the GR period as defined above. The last comparison period before the reform is 09/2007-08/2008, ending right before the GR period begins. Defining adjacent periods for before and during the GR can be justified given the sudden onset of the crisis with its sharp increase in unemployment (Figure 1).

Next, I systematically compare the GR with comparison periods after the GR, starting with 11/2011-10/2012 and shifting the comparison period until it covers 01/2014-12/2014. I do not start the comparison immediately after the period defined as the peak of the GR because the recession dragged on for a long time and unemployment decreased slowly (again, see Figure 1). I start in 11/2011 because using this comparison period, the difference in the entrepreneurial entry rate between the GR and the comparison period is significantly different from zero; when I use the comparison period beginning in 10/2011 or earlier, this difference is insignificant, which renders a decomposition obsolete.

For each decomposition, the figure shows the raw difference between the monthly entry rates into entrepreneurship in the GR and in the comparison period, the part of this difference explained by the observed individual characteristics, and the part of this difference explained by the distribution of the current unemployment status alone. It becomes very clear that the main results are robust to the choice of the comparison periods, especially when comparing with a period before the GR: With any time window chosen for comparison, almost the entire gap in the entrepreneurial entry rate is explained by the change in unemployment. Comparing with time windows after the GR is more difficult because of the slow recovery of the economy and employment. When choosing a comparison period close to the GR, the entrepreneurial entry rate is still almost as high as at the peak of the GR, and unemployment does not explain much of the difference because unemployment is also still almost as high as during the GR. Only later, when unemployment slowly falls, its contribution to the falling entry rate into entrepreneurship becomes more visible.

In summary, unemployment very robustly explains almost all of the increase in the entrepreneurial entry rate from before to during the GR. While unemployment explains a substantial part of the subsequent fall of the entry rate as well, much of this fall remains unexplained, especially during 2013. Therefore, an important avenue for further research

is to investigate which other factors became relevant after the GR that might account for this unexplained part of the decrease in the entry rate.[7]

5 Discussion and Conclusion

Entry rates into self-employment move with the unemployment rate over the business cycle. In this paper I use individual-level panel data from the matched monthly US Current Population Survey to decompose the change in the entrepreneurial entry rate into explained and unexplained components. The results indicate that the higher entry rate into entrepreneurship during the 2008/09 Great Recession is almost entirely explained by the higher unemployment rate during this period, together with a stable higher propensity of unemployed persons to become an entrepreneur in comparison to other individuals. Moreover, this counter-cyclicality of entrepreneurial entry out of unemployment is mostly driven by entry into unincorporated entrepreneurship, which is thus strongly related to necessity entrepreneurship (Fairlie and Fossen, 2018). In contrast, entry into incorporated entrepreneurship, which is related to opportunity entrepreneurship (Levine and Rubinstein, 2017; Herranz et al., 2017), is only weakly linked to the business cycle. These results contribute to consolidating seemingly contradictory findings from the extant literature. For example, Fairlie (2013) reports increasing entrepreneurship rates during the Great Recession in the United States, whereas Brunello and Langella (2016) find decreasing rates in Italy. The first study uses a broad measure of entrepreneurship including unincorporated businesses, whereas the second employs a narrower definition only including self-employed individuals who work as managers, professionals or in other skilled jobs, which is closer to our measure of incorporated entrepreneurship (see Shane, 2014).

The result that individual unemployment increases the probability of entry into en-

[7]Decker et al. (2016) report that startup rates rose during the 1990s and declined after 2000 in the high-tech sector. They also observe a generally declining dynamism after 2000.

trepreneurship is consistent with the "unemployment-push" hypothesis and in line with results from cross-sectional studies. By using individual-level panel data and econometric decomposition techniques, this study extends this research in an important direction by showing that unemployment largely explains the counter-cyclicality of entrepreneurship. Fairlie (2013) also uses panel data, but does not distinguish between unincorporated and incorporated entrepreneurship and does not decompose the entry probability.

The findings of this paper have important implications for the interpretation of the observed entry rate into entrepreneurship over the business cycle. The finding that unemployed individuals have a high probability of becoming entrepreneurs suggests that entrepreneurship enables workers who become unemployed during a recession to continue to use their human capital, which is likely to alleviate deterioration of human capital and unemployment scarring (Arulampalam et al., 2001). This way, entrepreneurship may facilitate a subsequent economic recovery, and public policy should therefore not impose barriers to this form of business formation. On the other hand, since the additional entry into entrepreneurship during a recession is mostly driven by unemployed individuals and since the additional firms formed are mostly unincorporated, one should expect lower levels of innovativeness and growth ambition of these entrepreneurs on average in comparison to start-ups during boom periods (cf. Ghatak et al., 2007). This result is complementary to that of Sedláček and Sterk (2017), who provide evidence that business cycle conditions at the time of firm formation matter for the performance of firm cohorts. These insights should be taken into account when forecasting the development of new businesses created during recessions and boom periods.

Further research should use data from other recessions and countries to assess the generality of the results found in this study. While the unemployment rate almost completely explains the difference between the entry rates before and during the Great Recession in the

United States, a larger part of the decline of the entrepreneurial entry rate after the Great Recession remains unexplained. An important avenue for further research is to find which additional factors related to individuals or the entrepreneurial ecosystem have recently emerged as relevant for entrepreneurial entry that were irrelevant before. In this context, more research is needed on the roles played by recent developments in new digital technologies that create novel opportunities in entrepreneurship or in automation technologies potentially pushing workers out of paid employment (Fossen and Sorgner, 2019).

References

Arulampalam, W., P. Gregg, and M. Gregory (2001). Unemployment scarring. *Economic Journal 111*(475), 577–584.

Åstebro, T., J. Chen, and P. Thompson (2011). Stars and misfits: Self-employment and labor market frictions. *Management Science 57*(11), 1999–2017.

Audretsch, D. B. and M. Vivarelli (1996). Determinants of new-firm startups in Italy. *Empirica 23*(1), 91–105.

Bartz, W. and A. Winkler (2016). Flexible or fragile? The growth performance of small and young businesses during the global financial crisis – Evidence from Germany. *Journal of Business Venturing 31*(2), 196–215.

Berglann, H., E. R. Moen, K. Røed, and J. F. Skogstrøm (2011). Entrepreneurship: Origins and returns. *Labour Economics 18*(2), 180–193.

Biehl, A. M., T. Gurley-Calvez, and B. Hill (2014). Self-employment of older Americans: do recessions matter? *Small Business Economics 42*(2), 297–309.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources 8*, 436–55.

Brunello, G. and M. Langella (2016). Local agglomeration, entrepreneurship and the 2008 recession: Evidence from Italian industrial districts. *Regional Science and Urban Economics 58*, 104–114.

Brünjes, J. and J. R. Diez (2013). 'Recession push' and 'prosperity pull' entrepreneurship in a rural developing context. *Entrepreneurship & Regional Development 25*(3-4), 251–271.

Caliendo, M., F. M. Fossen, A. Kritikos, and M. Wetter (2014). The gender gap in entrepreneurship: Not just a matter of personality. *CESifo Economic Studies 61*(1), 202–238.

Congregado, E., A. A. Golpe, and S. C. Parker (2012). The dynamics of entrepreneurship: hysteresis, business cycles and government policy. *Empirical Economics 43*(3), 1239–1261.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016). Where has all the skewness gone? The decline in high-growth (young) firms in the US. *European Economic Review 86*, 4–23.

Fairlie, R. W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement 30*(4), 305–316.

Fairlie, R. W. (2013). Entrepreneurship, economic conditions, and the Great Recession. *Journal of Economics & Management Strategy 22*(2), 207–231.

Fairlie, R. W. and F. M. Fossen (2018). Defining opportunity versus necessity entrepreneurship: Two components of business creation. Stanford SIEPR Working Paper 17-014.

Faria, J. R., J. C. Cuestas, and E. Mourelle (2010). Entrepreneurship and unemployment: a nonlinear bidirectional causality? *Economic Modelling 27*(5), 1282–1291.

Flood, S., M. King, S. Ruggles, and J. R. Warren (2017). Integrated Public Use Microdata Series, Current Population Survey: Version 5.0. Dataset, Minneapolis, MN: University of Minnesota.

Fortin, N., T. Lemieux, and S. Firpo (2011). Chapter 1 - decomposition methods in economics. Volume 4, Part A of *Handbook of Labor Economics*, pp. 1 – 102. Elsevier.

Fossen, F. M. and A. Sorgner (2019). New digital technologies and heterogeneous employment and wage dynamics in the United States: Evidence from individual-level data. IZA Discussion Papers 12242, Institute of Labor Economics.

Fritsch, M., A. Kritikos, and K. Pijnenburg (2015). Business cycles, unemployment and entrepreneurial entry – Evidence from Germany. *International Entrepreneurship and Management Journal 11*(2), 267–286.

Fritsch, M. and F. Noseleit (2013). Investigating the anatomy of the employment effect of new business formation. *Cambridge Journal of Economics 37*(2), 349–377.

Ghatak, M., M. Morelli, and T. Sjöström (2007). Entrepreneurial talent, occupational choice, and trickle up policies. *Journal of Economic Theory 137*(1), 27–48.

Henley, A. (2004). Self-employment status: The role of state dependence and initial circumstances. *Small Business Economics 22*(1), 67–82.

Herranz, N., S. Krasa, and A. P. Villamil (2017). Entrepreneurs, legal institutions and firm dynamics. *Economic Theory 63*(1), 263–285.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *Stata Journal 8*(4), 453–479.

Koellinger, P. D. and A. R. Thurik (2012). Entrepreneurship and the business cycle. *Review of Economics and Statistics 94*(4), 1143–1156.

Konon, A., M. Fritsch, and A. S. Kritikos (2018). Business cycles and start-ups across industries: An empirical analysis of German regions. *Journal of Business Venturing 33*(6), 742–761.

Lee, Y. and T. Mukoyama (2015). Entry and exit of manufacturing plants over the business cycle. *European Economic Review 77*, 20–27.

Levine, R. and Y. Rubinstein (2017). Smart and illicit: who becomes an entrepreneur and do they earn more? *Quarterly Journal of Economics 132*(2), 963–1018.

Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources 23*(3), 279–295.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review 14*, 693–709.

Oaxaca, R. L. and M. R. Ransom (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics 61*(1), 5–21.

Parker, S. C. (2018). *The economics of entrepreneurship*. Cambridge University Press.

Parker, S. C., E. Congregado, and A. A. Golpe (2012). Is entrepreneurship a leading or lagging indicator of the business cycle? Evidence from UK self-employment data. *International Small Business Journal 30*(7), 736–753.

Parker, S. C. and M. T. Robson (2004). Explaining international variations in self-employment: evidence from a panel of OECD countries. *Southern Economic Journal 71*(2), 287–301.

Ritsilä, J. and H. Tervo (2002). Effects of unemployment on new firm formation: Micro-level panel data evidence from Finland. *Small Business Economics 19*(1), 31–40.

Sedláček, P. and V. Sterk (2017). The growth potential of startups over the business cycle. *American Economic Review 107*(10), 3182–3210.

Shane, S. (2014). Incorporated versus unincorporated self employment. Small Business Trends, https://smallbiztrends.com/2013/11/incorporated-versusunincorporated-self-employment.html.

Storey, D. J. (1991). The birth of new firms–does unemployment matter? A review of the evidence. *Small Business Economics 3*(3), 167–178.

Thurik, A. R., M. A. Carree, A. Van Stel, and D. B. Audretsch (2008). Does self-employment reduce unemployment? *Journal of Business Venturing 23*(6), 673–686.

Yun, M.-S. (2004). Decomposing differences in the first moment. *Economics Letters 82*(2), 275–280.

Yun, M.-S. (2005). Normalized equation and decomposition analysis: Computation and inference. IZA Discussion Papers 1822, Institute of Labor Economics.

Tables

Table 1: Means and standard deviations before, during and after the Great Recession

	Before Great Recession		Great Recession		After Great Recession	
Variable	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Entry into entrep.	0.0054	0.0734	0.0061	0.0780	0.0057	0.0751
Entry into uninc. entrep.	0.0043	0.0654	0.0048	0.0695	0.0044	0.0661
Entry into inc. entrep.	0.0011	0.0334	0.0013	0.0357	0.0013	0.0359
Unemployed	0.0370	0.1889	0.0663	0.2488	0.0545	0.2270
Male	0.4703	0.4991	0.4723	0.4992	0.4757	0.4994
Less then high school	0.0924	0.2896	0.0897	0.2858	0.0800	0.2713
High school degree	0.2926	0.4550	0.2913	0.4544	0.2720	0.4450
Some college	0.2983	0.4575	0.2985	0.4576	0.3010	0.4587
University degree	0.3166	0.4652	0.3205	0.4667	0.3470	0.4760
Age	40.8053	11.7057	40.9806	11.8066	41.1586	12.1260
White	0.8270	0.3782	0.8258	0.3793	0.8116	0.3910
African American	0.0957	0.2942	0.0953	0.2936	0.1004	0.3006
Other nonwhite	0.0773	0.2671	0.0789	0.2696	0.0880	0.2833
Married	0.6089	0.4880	0.6045	0.4890	0.5769	0.4941
Number of children	0.9707	1.1855	0.9660	1.1876	0.9530	1.1896
Metropolitan area	0.8003	0.3997	0.8009	0.3993	0.8148	0.3884
West	0.2515	0.4339	0.2486	0.4322	0.2543	0.4355
Northeast	0.2027	0.4020	0.2052	0.4039	0.1950	0.3962
Midwest	0.2378	0.4258	0.2405	0.4274	0.2352	0.4241
South	0.3080	0.4617	0.3057	0.4607	0.3155	0.4647
Observations	538469		541517		518355	

Note: The period before the Great Recession is 04/2007-03/2008, during the GR 09/2008-08/2009, and after the GR 05/2013-04/2014. *Source:* Own calculations based on the Current Population Survey.

Table 2: Probability of entry into entrepreneurship: Marginal effects from logit estimations

	Before Great Recession	Great Recession	After Great Recession
Unemployed	0.00928***	0.00940***	0.00868***
	(0.000324)	(0.000302)	(0.000313)
Male	0.00151***	0.00157***	0.00124***
	(0.000201)	(0.000214)	(0.000211)
High School	-0.000678*	-0.00198***	-0.00164***
	(0.000346)	(0.000353)	(0.000369)
Some college	-0.00111***	-0.00210***	-0.00159***
	(0.000354)	(0.000359)	(0.000368)
University	-0.000281	-0.00147***	-0.00124***
	(0.000348)	(0.000356)	(0.000363)
Black	-0.00109***	-0.00200***	-0.00180***
	(0.000382)	(0.000418)	(0.000397)
Oth. nonwhite	-0.000923**	-0.000711*	-0.000346
	(0.000396)	(0.000415)	(0.000375)
Married	0.000428*	0.000556**	0.0000266
	(0.000236)	(0.000250)	(0.000243)
No. children	0.000289***	0.000170*	0.000287***
	(0.0000959)	(0.0000993)	(0.0000991)
Metropolitan	-0.000724***	-0.000508*	-0.000473*
	(0.000242)	(0.000262)	(0.000268)
Age and age sq.	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes
Average probability	0.00541	0.00612	0.00568
Observations	538469	541517	518355

Notes: Logit estimations of the monthly probability of entry into entrepreneurship before, during and after the Great Recession. The period before the GR is 04/2007-03/2008, during the GR 09/2008-08/2009, and after the GR 05/2013-04/2014. Average marginal effects. Robust standard errors in parentheses. *,**,***: Significant at the 10%/5%/1%-levels. *Source:* Own calculations based on the Current Population Survey.

Table 3: Probability of entry into (un)incorporated entrepreneurship: Marginal effects from logit estimations

	Unincorporated entrepreneurship			Incorporated entrepreneurship		
	Before GR	GR	After GR	Before GR	GR	After GR
Unemployed	0.00781***	0.00790***	0.00718***	0.00123***	0.00135***	0.00131***
	(0.000284)	(0.000268)	(0.000273)	(0.000177)	(0.000151)	(0.000166)
Male	0.000978***	0.000967***	0.000461**	0.000512***	0.000586***	0.000771***
	(0.000178)	(0.000190)	(0.000184)	(0.0000946)	(0.000101)	(0.000108)
High School	-0.00101***	-0.00217***	-0.00156***	0.000886***	0.000885***	0.000113
	(0.000290)	(0.000297)	(0.000308)	(0.000260)	(0.000265)	(0.000229)
Some college	-0.00165***	-0.00249***	-0.00169***	0.00111***	0.00110***	0.000315
	(0.000302)	(0.000306)	(0.000310)	(0.000259)	(0.000265)	(0.000224)
University	-0.00129***	-0.00241***	-0.00191***	0.00149***	0.00152***	0.000774***
	(0.000299)	(0.000309)	(0.000312)	(0.000259)	(0.000263)	(0.000218)
Black	-0.000951***	-0.00167***	-0.00154***	-0.000131	-0.000324*	-0.000256
	(0.000342)	(0.000371)	(0.000350)	(0.000171)	(0.000194)	(0.000187)
Oth. nonwhite	-0.000922**	-0.000531	-0.000454	-0.0000235	-0.000181	0.0000917
	(0.000358)	(0.000368)	(0.000334)	(0.000172)	(0.000194)	(0.000172)
Married	0.000112	0.000114	-0.000357*	0.000341***	0.000479***	0.000421***
	(0.000209)	(0.000220)	(0.000213)	(0.000115)	(0.000125)	(0.000123)
No. children	0.000184**	0.000136	0.000184**	0.0000946**	0.0000212	0.0000863*
	(0.0000868)	(0.0000889)	(0.0000870)	(0.0000417)	(0.0000447)	(0.0000481)
Metropolitan	-0.001000***	-0.000679***	-0.000814***	0.000367***	0.000206	0.000432***
	(0.000210)	(0.000230)	(0.000228)	(0.000131)	(0.000131)	(0.000151)
Age & age sq.	Yes	Yes	Yes	Yes	Yes	Yes
Regional dum.	Yes	Yes	Yes	Yes	Yes	Yes
Average probability	0.00430	0.00485	0.00439	0.00112	0.00127	0.00129
Observations	538469	541517	518355	538469	541517	518355

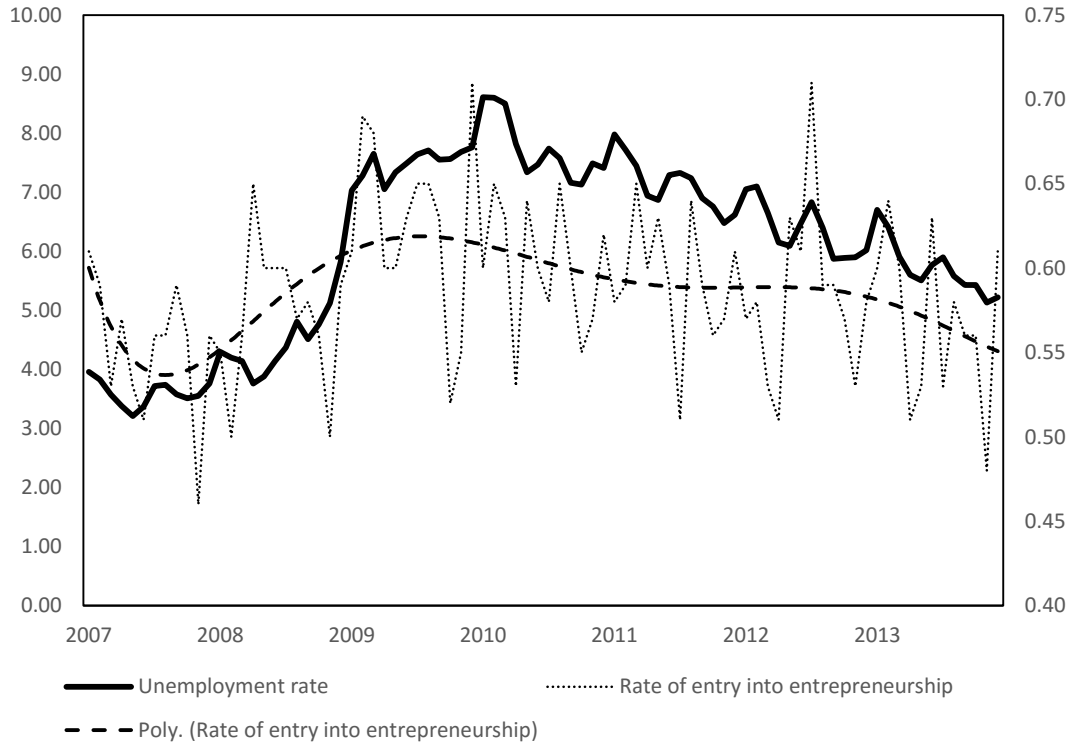
Notes: Logit estimations of the monthly probability of entry into unincorporated and incorporated entrepreneurship before, during and after the Great Recession. The period before the GR is 04/2007-03/2008, during the GR 09/2008-08/2009, and after the GR 05/2013-04/2014. Average marginal effects. Robust standard errors in parentheses. *,**,***: Significant at the 10%/5%/1%-levels. *Source:* Own calculations based on the Current Population Survey.

Table 4: Oaxaca nonlinear decomposition of entry rate into entrepreneurship

	All entrepreneurship		Unincorp. entrepreneurship		Incorp. entrepreneurship	
	GR versus before	GR vs. after	GR vs. before	GR vs. after	GR vs. before	GR vs. after
Overall difference in entrepreneurial entry rates						
Great Rec.	0.00612***	0.00612***	0.00485***	0.00485***	0.00127***	0.00127***
	(0.000106)	(0.000106)	(0.0000942)	(0.0000942)	(0.0000485)	(0.0000485)
Comp. period	0.00541***	0.00568***	0.00430***	0.00439***	0.00112***	0.00129***
	(0.0000999)	(0.000104)	(0.0000890)	(0.0000917)	(0.0000455)	(0.0000499)
Difference	0.000708***	0.000442***	0.000552***	0.000461***	0.000156**	-0.0000184
	(0.000146)	(0.000149)	(0.000130)	(0.000131)	(0.0000665)	(0.0000696)
Explained	0.000571***	0.000242***	0.000505***	0.000237***	0.0000619***	0.00000596
	(0.0000208)	(0.0000139)	(0.0000192)	(0.0000125)	(0.00000745)	(0.00000538)
Unexplained	0.000138	0.000200	0.0000465	0.000223*	0.0000943	-0.0000243
	(0.000145)	(0.000148)	(0.000129)	(0.000131)	(0.0000661)	(0.0000693)
Difference explained by individual characteristics						
Unemployed	0.000567***	0.000197***	0.000511***	0.000163***	0.0000540***	-0.000979
	(0.0000209)	(0.0000110)	(0.0000196)	(0.00000902)	(0.00000695)	(0.0522)
Male	0.00000654**	-0.00000871***	0.00000443**	-0.00000440***	0.00000160**	0.000142
	(0.00000312)	(0.00000277)	(0.00000216)	(0.00000155)	(0.000000774)	(0.00759)
Education	-0.00000410	0.0000108	-0.0000122***	0.0000490***	0.00000687***	0.00140
	(0.00000253)	(0.00000925)	(0.00000296)	(0.00000784)	(0.00000135)	(0.0745)
Race	-0.00000148	0.0000268***	-0.00000145	0.0000234***	-0.000000105	-0.000113
	(0.00000218)	(0.00000580)	(0.00000202)	(0.00000512)	(0.000000357)	(0.00604)
Married	-0.00000449**	0.0000155*	-0.00000109	-0.00000540	-0.00000255***	-0.000774
	(0.00000186)	(0.00000860)	(0.00000151)	(0.00000779)	(0.000000792)	(0.0414)
No. children	-0.00000226*	0.00000542***	-0.00000170	0.00000376**	-0.000000390	-0.0000440
	(0.00000130)	(0.00000196)	(0.00000105)	(0.00000164)	(0.000000278)	(0.00234)
Metropolitan	-0.000000701	0.0000127***	-0.00000103	0.0000190***	0.000000218	0.000271
	(0.00000102)	(0.00000482)	(0.00000147)	(0.00000422)	(0.000000317)	(0.0144)
Age	0.0000238***	0.0000195**	0.0000183***	0.0000151*	0.00000450***	-0.000225
	(0.00000601)	(0.00000975)	(0.00000502)	(0.00000820)	(0.00000155)	(0.0121)
Region	-0.0000136***	-0.0000364***	-0.0000111***	-0.0000263***	-0.00000222***	0.000330
	(0.00000359)	(0.00000611)	(0.00000314)	(0.00000510)	(0.000000749)	(0.0176)
N	1079986	1059872	1079986	1059872	1079986	1059872
N: GR	541517	541517	541517	541517	541517	541517
N: Comp. per.	538469	518355	538469	518355	538469	518355

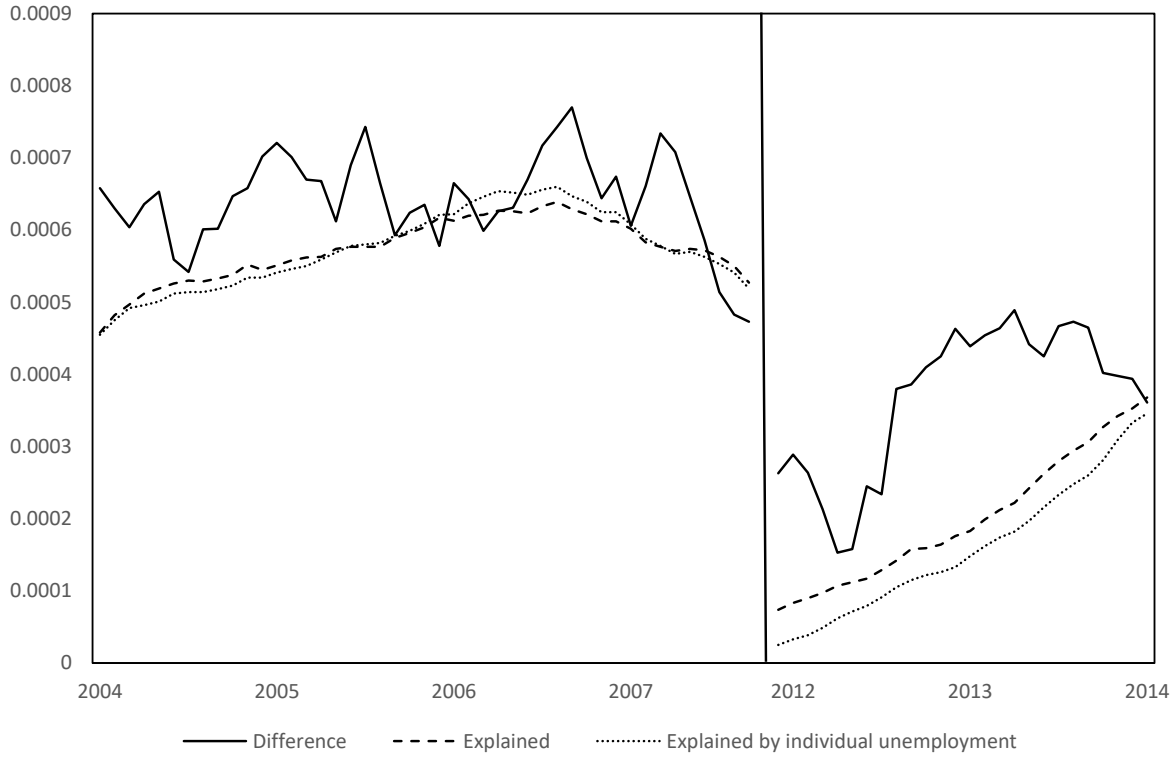
Notes: Nonlinear Oaxaca decomposition of the monthly entry rate into entrepreneurship based on logit estimations. The columns show separate decompositions of the difference in the entry rate during the Great Recession and a comparison period before or after the Great Recession. Separate decompositions for entry into all, unincorporated and incorporated entrepreneurship are shown. The lower panel shows the difference between the entry rates in the two periods explained by each of the variables. The period before the GR is 04/2007-03/2008, during the GR 09/2008-08/2009, and after the GR 05/2013-04/2014. Robust standard errors in parentheses. *,**,***: Significant at the 10%/5%/1%-levels. *Source:* Own calculations based on the Current Population Survey 2007-2014.

Figures



Note: Unemployment rate (left scale) and entry rate into entrepreneurship (right scale) in the United States in %. The dashed line shows a sixth degree polynomial fit of the entry rate. *Source:* Own illustration based on the Current Population Survey 2007-2013.

Figure 1: Unemployment and Entrepreneurial Entry



Note: The figure shows the results of the decomposition of the change in the entrepreneurial entry rate when comparing a shifting time window to the Great Recession period (9/2008-8/2009). The comparison period is always a year wide. For each month in the graph, the results for the comparison period starting in that month are displayed. The solid line shows the difference between the entrepreneurial entry rates in the Great Recession and the comparison period, the dashed line the explained part of this difference, and the dotted line the part of the difference explained by the unemployment rate. For example, for 04/2007 and 05/2013, the figure shows the same results as Table 4. No results are shown for 10/2007-10/2011 (solid vertical line) when the comparison period would partially overlap with the recession. *Source:* Own illustration based on the Current Population Survey 2004-2014.

Figure 2: Decomposition Results with Shifting Comparison Period

Appendix

The logit model of the probability of entry into entrepreneurship can be written as:

$$Y = F(X\beta) \tag{A.1}$$

where Y is the vector of predicted entry probabilities, X the matrix of independent variables, β the coefficient vector, and F the cumulative logistic distribution function. A non-linear decomposition of the mean difference in entrepreneurial entry by period can be written as:

$$\overline{Y}_R - \overline{Y}_C = \left[\overline{F(X_R\beta_R)} - \overline{F(X_C\beta_R)}\right] + \left[\overline{F(X_C\beta_R)} - \overline{F(X_C\beta_C)}\right] \tag{A.2}$$

where index R stands for the observations during the Great Recession and index C for the observations in the comparison period. In Equation A.2 the first summand is the contribution of the distribution of the variables to the overall difference in the entry rate, i.e. the explained part, whereas the second summand is the contribution of differences in the coefficients (including the constant), i.e. the unexplained part.[8] Following the approach of Yun (2004), for a detailed decomposition which assesses the contributions of each single variable (or group of variables) separately in this non-linear setting, two approximations are necessary. First, I consider predictions at the mean values of the explanatory variables:

$$\overline{Y}_R - \overline{Y}_C = \left[F(\overline{X}_R\beta_R) - F(\overline{X}_C\beta_R)\right] + \left[F(\overline{X}_C\beta_R) - F(\overline{X}_C\beta_C)\right] + R_A, \tag{A.3}$$

where

$$\begin{aligned} R_A = &\left[\overline{F(X_R\beta_R)} - \overline{F(X_C\beta_R)}\right] + \left[\overline{F(X_C\beta_R)} - \overline{F(X_C\beta_C)}\right] \\ &- \left[F(\overline{X}_R\beta_R) - F(\overline{X}_C\beta_R)\right] - \left[F(\overline{X}_C\beta_R) - F(\overline{X}_C\beta_C)\right]. \end{aligned} \tag{A.4}$$

[8] More precisely, as mentioned in Section 2, we use the coefficient estimates from a pooled estimation for the decomposition of the contributions of the observed characteristics to the differential; see Jann (2008) for the technical details.

Second, a first order Taylor expansion around the mean characteristics is used. Hence, I can rewrite Equation A.3 as follows:

$$\begin{aligned} \overline{Y}_R - \overline{Y}_C &= \left[(\overline{X}_R - \overline{X}_C)\beta_R\right] f(\overline{X}_R\beta_R) \\ &+ \overline{X}_C(\beta_R - \beta_C) f(\overline{X}_C\beta_C) + R_A + R_T, \end{aligned} \tag{A.5}$$

where $f(\cdot)$ is the first order derivative of $F(\cdot)$ and R_T is the approximation error. Using A.5, a detailed decomposition of Equation A.2 can be written as

$$\overline{Y}_R - \overline{Y}_C = \sum_{i=1}^{K} W^i_{\Delta X} \left[\overline{F(X_R\beta_R)} - \overline{F(X_C\beta_R)}\right] + \sum_{i=1}^{K} W^i_{\Delta\beta} \left[\overline{F(X_C\beta_R)} - \overline{F(X_C\beta_C)}\right], \tag{A.6}$$

i.e., the detailed decomposition includes weights for the contributions of the characteristics ($W^i_{\Delta X}$) and for the contributions of the coefficients ($W^i_{\Delta\beta}$), with

$$W^i_{\Delta X} = \frac{(\overline{X}^i_R - \overline{X}^i_C)\beta^i_R}{(\overline{X}_R - \overline{X}_C)\beta_R} \quad \text{and} \quad W^i_{\Delta\beta} = \frac{\overline{X}^i_C(\beta^i_R - \beta^i_C)}{\overline{X}_C(\beta_R - \beta_C)}$$

for variable i in the set of K explanatory variables (Yun, 2004).