Knockin' on the Bank's Door: Why is Self-Employment Going Down?

Alina Malkova*

January 2, 2020

Abstract

This study examines the decline in the ability to obtain financing as a potential explanation for the observed decrease in U.S. self-employment. The shrinking of the U.S. branch bank network since 2010 and increased average borrower-lender distance decreases the accessibility of the credit institutions for borrowers. To evaluate the impact of the credit market accessibility (CMA) on entry into self-employment, I disaggregate the self-employed into two categories: entrepreneurs whose businesses depend on business loans (Type-1) and other self-employed (Type-2). Using a novel data source (the Community Advantage Panel Survey database), I find that the proximity of credit market institutions has heterogeneous effects on the transition to self-employment. An improvement in CMA increases the likelihood of transition to Type-1 self-employment. But, for Type-2 self-employment, the effect is the opposite: the transition probability to Type-2 self-employment decreases, and this type of self-employed more likely switches to paid-employment to be able to receive non-business related loans. Based on the estimates, the paper determines the effect of several improvements in credit market access (e.g., additional branch banks) on the two types of self-employment and paid-employment.

*Department of Economics, University of North Carolina at Chapel Hill. E-mail: amalkova@live.unc.edu.
I would like to thank Klara Peter, Luca Flabbi, David Guilkey, Andres Hincapie, and Helen Tauchen for serving on my thesis committee and providing valuable guidance and feedback. I am grateful for discussions with many people, including Donna Gilleskie, Jane Fruehwirth, Qing Gong, Siddhartha Biswas, Mauricio Salazar, Ray Wang, and participants of the Applied Micro seminar at the University of North Carolina at Chapel Hill. I deeply appreciate Andrey Minaev for helpful comments and support. I gratefully acknowledge the Center of Community Capital UNC, the Ford Foundation and Sarah Riley for providing access to the CAPS database, helpful feedback, and technical support. Any errors are my own.

1 Introduction

The self-employed sector in the U.S. has lost almost 1 million workers since 1994[1] (see Fig.1). The share of self-employed in total employment decreased from 12.1 percent in 1994 to 10.1 percent in 2015 (Hipple, 2016)[2]. The literature provides limited reasoning for the decrease in the number of self-employed. According to Hipple (2016), the decline in self-employment may be associated with a decline in agricultural employment. The other reason for this decline is documented by Schweitzer and Shane (2016). Authors find that business cycles affect transitions in and out of self-employment,[3] while the recent fall in the aggregate demand explains the exit from entrepreneurship. This paper focuses on the decline in the ability to obtain financing as another potential explanation for the observed trends in self-employment.

In recent years, two factors may have contributed to the reduction of credit market accessibility among the self-employed. The de-liberalization of the financial system during the Great Recession increased regulatory scrutiny and tightened collateral requirements for borrowers (Wiersch and Shane, 2013). But the recovery of the financial system and the easing of requirements for borrowers did not lead to pre-crisis indicators of self-employment. The other factor is the shrinking of the U.S. branch network since 2010 and, as a consequence, the increased average borrower-lender distance (see Fig.3) (Nguyen, 2019). The Federal Deposit Insurance Company (FDIC) claims that at least 80 percent of branch closings should not have any meaningful impact on physical access to banks, as they occur in the areas with multiple remaining branches.[4] However, the borrowing process for the self-employed is informationally-intensive,

[1] The number is reported by the U.S. Bureau of Labor Statistics (number of unincorporated self-employed persons). The definition of self-employed individuals comes from the Bureau of Labor Statistics (BLS). Self-employed people are those who work for profit or fees in their own business, profession, trade, or farm, including those who intended to earn a profit but whose business produced no profit or a loss. Unless otherwise specified, CPS estimates of the self-employed published by BLS reflect only people whose businesses are unincorporated. BLS classifieds the incorporated self-employed as wage and salary workers, because the incorporated self-employed are paid employees of their corporation.

[2] From 1994 to 2015, the unincorporated self-employment rate fell from 8.7 percent in 1994 to 6.4 percent in 2015. From 1994 to 1999, the share of total employment made up of the incorporated self-employed ranged from 3.2 percent to 3.5 percent. Over the 2000–08 period, the incorporated self-employment rate rose from 3.3 percent to 4.0 percent. The rate then edged down to 3.7 percent in 2010 and remained at that level over the 2011–15 period (see Fig.2).

[3] Beckhusen (2014) also shows that the transitions from self-employment to wage-work increased in the post-recession months.

[4] The closing of branches is heavily regulated by the FDIC (`https://www.fdic.gov/regulations/laws/rules/5000-3830.html`). Before closing a branch, banks must provide detailed research and statistics about

and credit approvals primarily rely on soft information about the borrower (DeYoung et al., 2008). In other words, personal presence during a loan transaction can still play an essential role in obtaining financing.

This study confirms that the credit market accessibility has substantially declined in the U.S. from 2007 to 2014 (see Fig.4), where the accessibility is proxied by the average distance between a borrower and a lender, the number of nearby branches, and the number of bank employees in the county of residence. However, the effect of the reduced physical credit market accessibility on self-employment rates is ambiguous. On the one hand, the relationship between the credit market accessibility (CMA) and the probability of being self-employed could be positive. A few previous studies find that a shorter lender-borrower distance increases business lending (Nguyen (2019); Degryse and Ongena (2005)), which may positively influence the number of entrepreneurs. On the other hand, studies from countries with a large informal sector show that new bank openings create incentives for individuals working in the informal sector to switch to jobs with verifiable income in order to be eligible for the loan and thus transfer to salary employment (Malkova et al., 2019). In the case of the USA, the decline in the borrower-lender distance may create incentives for self-employed workers with restricted access to the credit market[5] to switch to paid-employment so that they may be able to use credit products. In other words, the effect of CMA on the transition to self-employment is likely to be heterogeneous, which is the focus of this study.

The study utilizes several sources of U.S. data between 2003 and 2014. The main database is the Community Advantage Panel Survey (CAPS)[6] database. The advantage of using this database is the possibility to identify geographical distances between each individual and all banks within the state. The location of banks is taken from the FDIC (SOD) data. To diminish the effect of unobservable local economic conditions that are correlated with the credit supply, I also use county and tract-level control variables from the American Community Survey (ACS).

the effect of the closure on the local economy, and clients must receive a notice 90 days before the branch closes in the given area.

[5] For example, Johnson (2016) stays that due to the unstable, and sometimes informal, nature of business income of self-employed means that the business income is often not counted in full when self-employed applies for a mortgage and lenders calculate the debt-to-income ratio. The author shows that the Ability-to-Repay rule significantly restricts self-employed households' access to mortgage credit. By restricting credit access, verified debt-to-income requirements lead to a reduction in self-employment and small business employment.

[6] CAPS comprises 11 years of a panel survey of approximately 5,000 low- and moderate-income homeowners and renters during the period of 2003-2014. The data collection has been conducted by the UNC Center for Community Capital with generous funding from the Ford Foundation.

The paper proceeds in two parts. In the first part, I present testable implications from an intuitive theoretical model (a three-sector Roy model) that shows the effect of credit market accessibility on labor mobility among three labor states: two types of self-employment (SE) and paid-employment (wage employment). The theoretical model is based on the work of Evans and Jovanovic (1989) and Levine and Rubinstein (2018). To show the heterogeneous effects of CMA on the decision to become self-employed, I divide self-employed workers into two types. The first category of self-employed (or Type-1 self-employed) are entrepreneurs, owners of firms that demand physical capital and business loans[7]. The Type-2 self-employed are non-entrepreneurs and owners of companies without business loans; in other words, the agent uses own initial wealth to start a business[8]. In the theoretical model, agents enter self-employment (Type-1 and Type-2) based on their comparative advantage that depends on their abilities and the amount of available assets. Individuals face borrowing costs including both the interest costs and non-interest costs (the time costs of arranging a loan and any monetary initiation costs). The comparative static shows that smaller non-interest costs of borrowing (better credit market accessibility) increases the probability of being Type-1 self-employed, decreases the probability of being Type-2 self-employed, and increases the transition probability from Type-2 self-employment to paid-employment.

The second part of the paper is the empirical analysis that aims to estimate the impact of credit market accessibility on the transition in and out of self-employment. The CMA is measured as an index based on the combination of the following variables: average distance from the individual to the 10 nearest banks, the number of banks and branches within 5 miles from the individual, the number of bank workers per 1000 pop. at the county level. The model used in the analysis is the dynamic multinomial logit model of employment with correlated random effects. The empirical challenge is that banks choose where to locate branches depending on local economic conditions that are correlated with the number of self-employed in the given area. To identify the causal impact of CMA on the decision to become self-employed, I use several approaches. First, I control for the set of variables that proxy local economic conditions at the county level: the number of business establishments, unemployment rate, the share of the population with professional, scientific, management and ad- min. education. Second, I include in the model the within-means of CMA characteristics that accounts for the endogeneity of CMA

[7] As an example of Type-1 SE, mention may be made of computer programmers, lawyers, doctors, real estate managers, etc.

[8] Examples of Type-2 self-employed are tutors, housekeeping cleaners, lawn service workers, grounds maintenance workers, etc.

with regards to unobserved factors like the fixed-effect model. Also I estimate the model using exposure to post-merger consolidation of banks at the Census tract-level as an instrument for changes in the credit market accessibility (Nguyen (2019); Garmaise and Moskowitz (2006)). In this paper, I consider mergers of large banks, those where both banks held at least $1 billion in pre-merger assets. The average Census tract is 1.5 square meters, and the decision of large banks to merge is plausibly exogenous to local economic conditions in Census tract where both banks have a branch. The merge-induced consolidation decreases banks' competition at the tract level, which may be followed by a branch closure (Nguyen (2019)) and the increase in the borrower-lender distance, decreasing in the number of branches employees and the decline in the local credit market accessibility. The empirical analysis shows heterogeneous effects of CMA on switching in and out of self-employment depending on the type of self-employed workers. For the Type-1 self-employed, the improvement of credit market accessibility increases the probability of transition to self-employment. While for the Type-2 self-employed, the improvement of credit-market accessibility increases the probability of transition out of self-employment.

This paper proceeds as follows. Section 2 contains a literature review. Section 3 discusses testable implications from the theoretical model. Section 4 develops the empirical model. Section 5 discusses data. Section 6 presents the results and considers policy simulations.

2 Literature Review

The literature review focuses on two segments of the extensive and growing literature on entrepreneurship and self-employment (Parker, 2018). The first are factors that influence the decision to become self-employed. The second is the impact of credit market accessibility, particularly geographic proximity on self-employment.

One strand of the literature examines how the decision to become an entrepreneur is associated with access to capital, wealth, and collateral constraints. A large amount of literature focuses on the positive correlation between housing wealth and entrepreneurial activities (Fan and White (2003); Fairlie and Krashinsky (2012); Fort et al. (2013); Corradin and Popov (2015)). However, Hurst and Lusardi (2004) show that the propensity to become a business owner is a nonlinear function of wealth. They show that the relationship between wealth and entry into entrepreneurship is mostly flat over the majority of the wealth distribution. It is only at the top of the wealth distribution (greater than the 95th percentile) that a positive relationship can be found.

The literature shows that credit constraints at the household level matter for the creation of new businesses (Evans and Jovanovic (1989); Holtz-Eakin et al. (1994); Gentry and Hubbard (2004); Cagetti and De Nardi (2006)). Previous work has also found that a bank loan is an essential source of financing for small businesses (Petersen and Rajan (1994); Bates and Robb (2013); Fracassi et al. (2013)) and that entrepreneurs often have to provide personal guarantees when they obtain financing (Berger and Udell (1998); Greenstone et al. (2014)). Herkenhoff et al. (2016) examine how consumer credit access impacts employment prospects, earnings, and entrepreneurship. To isolate the causal impact of credit on labor market outcomes, authors use bankruptcy flag removals[9] to separate a sizable discrete increase in credit access. Authors show that following the flag removal there is (a) an increased flow rate into self- employment, (b) disproportionate borrowing by new self-employed entrants relative to other job transitioners, (c) an increased likelihood of starting an employer business, (d) startups enter capital intensive and external finance intensive industries, and (e) disproportionate borrowing by new employer businesses.

Geographic distance still matters for bank lending. Two broad channels have been identified in economics literature for the effects of distance on credit market transactions. Firstly, studies on spatial rationing have established a correspondence between distance and credit rationing. A closer geographic distance gives banks privileged access to soft information that facilitates the evaluation of the borrower's creditworthiness, thereby permitting them to gain a cost advantage for monitoring over more remote competitors who may not enjoy the same degree of access to such information (Hauswald and Marquez, 2006). The information effect of distance has been shown empirically to facilitate ex ante screening and ex post monitoring of borrowers in bank lending, providing well-informed banks with competitiveness and market power (Petersen and Rajan (2002); Brevoort et al. (2010); Guiso et al. (2004); Sufi (2007); Qian and Strahan (2007)). One benefit for borrowers closer to their banks is that, with privileged access to information, inefficient rationing might be reduced. Using a unique data set of all loan applications by small firms to a large bank, Agarwal and Hauswald (2010) document that the closer a firm is to its branch office, the more likely the bank is to offer credit. However, borrowing from closer banks is associated with, for example, higher interest rates (Agarwal and Hauswald, 2010). The reason is that borrowers are informationally captured by lenders who have privileged access to soft

[9] A bankruptcy filing raises a "red flag" to potential lenders and will affect future lending opportunities. It will likely prevent an individual from obtaining credit for a considerable period of time after filing. Even when a debtor is able to obtain credit, the associated interest rate will be exorbitant. Under the Fair Credit Reporting Act, a bankruptcy filing can remain on an individual's credit report for 10 years, and the bankruptcy can be removed from the credit report after this term.

information to the extent that such information cannot be credibly communicated to outsiders (Dell'Ariccia and Marquez, 2004).

Secondly, shorter lender-borrower distance potentially benefits both parties since it reduces transaction costs. Examples of such costs for the borrower include transportation costs incurred in the process of the loan application and the time and effort spent by a potential borrower to either personally interact with loan officers or to look for a suitable loan. The reduction in the cost of obtaining soft information for the lender is another benefit of the closer distance. In classical models of location differentiation (Salop, 1979), borrowers incur distance-related transportation costs from visiting their banks while the banks price loans uniformly if they cannot observe borrowers' locations or are prevented from charging different prices to different borrowers (Freixas and Rochet, 2008). However, if banks observe the borrowers' locations and offer interest rates based on that information, they may engage in spatial price discrimination (Lederer and Hurter Jr, 1986). For example, Degryse and Ongena (2005) document that loan rates decrease with the distance between the firm and lending bank and increase with the distance between the firm and competing banks, suggesting that transportation costs cause the spatial price discrimination. Nguyen (2019) shows that the closing of bank branches since 2005 lead to a persistent decline in local small business lending. The author shows that the effect is very localized and dissipate within 6 miles from the borrower's location.

Contribution to the Literature

This paper contributes to the literature on self-employment in many ways. First, the literature has not explored the role of the credit market channel in self-employment trends broadly. As it was mentioned above, the literature discusses the importance of geographical proximity for business lending, but not in the context of very small borrowers such as self-employed. In this paper, I use a novel data source that allows for the identification of geographical distances between respondents' homes and all banks within a state. To the best of my knowledge, this paper is the first to study the role of a borrower-lender distance in the decision to become self-employed.

Second, the paper contributes to the literature discussing the diversity of the self-employed. A growing literature argues that the self-employed should not be considered as an aggregated group but instead split into categories with several different groupings suggested. Levine and Rubinstein (2017) show that there are crucial differences between unincorporated and incorporated businesses. Block and Sandner (2009) argue that in using push and pull factors which determine the selection into self-employment, we can divide the self-employed into two groups: "necessity" entrepreneurs who self-employ due to the lack of other options and "opportunity" entrepreneurs who seek to bring new ideas to the market or avail themselves of other market advantages. This paper provides a framework in which individuals split into two types based on the dependence of their enterprises on outside financing. The Type-1 self-employed are business owners of firms that demand physical capital and business loans. The Type-2 self-employed are owners of businesses that demand less capital and without business loans; in other words, the agent uses her initial wealth to start a business.

Third, the paper shows the heterogeneous effect of CMA on the transition to self-employment. The improvement in credit market accessibility increases the entry into entrepreneurship for those individuals whose business depends on business loans, but it also increases the transition from self-employment to paid-employment if the individual has restricted access to the credit market being self-employed.

3 Testable Implications from the Theoretical Model

This section describes the highlights of the theoretical model - an intuitive three-sector Roy model of labor market decisions that (a) captures the essential features of the relationship between the access to the credit market and labor mobility among three labor states: the two types of self-employment (SE_1 and SE_2) and paid employment (or salaried employment) (PE), (b) provides testable predictions for the empirical work.

The theoretical model is based on the works of Evans and Jovanovic (1989) and Levine and Rubinstein (2018). I divide self-employed workers into two categories. The first category of self-employed (or Type-1 self-employed) are owners of businesses that demand entrepreneurial skills, physical capital, and business loans. As an example of this category of self-employed, mention may be made of computer programmers, lawyers, doctors, real estate managers, etc. In contrast with Levine and Rubinstein (2018), who state that the second category of self-employed (or Type-2 self-employed) demand "none or little entrepreneurial ability, physical capital and liquidity" and are driven mostly by non-pecuniary benefits of self-employment, I use another definition for this type of self-employment and model the Type-2 self-employed as non-entrepreneurs and owners of businesses without business loans; in other words, either the agent does not have enough collateral to access the credit market, or the agent can use her initial wealth to start a business (if her business is not capital intensive). As an example of this category of self-employed, mention may be made of tutors, babysitters, maintenance workers, etc.

The model is a discrete-time model. In each period, individuals select among the three labor states, where changes between states include known switching costs. Given the information available at the beginning of each period, individuals choose their labor states, borrowings, and consumption to maximize discounted expected utility. Individuals face borrowing costs including both the interest costs and non-interest costs ψ (the time costs of arranging a loan and any monetary initiation costs)[10].

The mechanisms for the effects of credit accessibility are through the market for loans. As is perhaps obvious, improved credit market access makes large-scale, incorporated (Type-1) self-employment more attractive since such business opportunities often require outside financing. It is perhaps not as obvious that improved credit market conditions may also make paid employment more attractive in comparison to small-scale, informal (Type-2) self-employment. Many lenders require demonstrable evidence of stable income for consumer loans (e.g., auto, mortgage, credit cards), and providing such evidence may be easier with paid employment than small-scale, informal self-employment. Thus, as credit market conditions improve, and individuals are more likely to seek loans, they may also seek employment that allows them to satisfy the income verification conditions for the loans. This phenomenon has been considered previously for developing countries, but very little study has been done for developing countries.

As shown in the Appendix, the theoretical model yields testable implications on the impact of available assets and credit market accessibility on labor market states. The key testable implication of the model is that an increase in credit market access attributable to a reduction in the non-interest costs of borrowing, ψ_t and ψ_{t+1}, increasing the probability of being Type-1 self-employed. Also, the theoretical model shows that the likelihood of being Type-1 self-employed increases if the agent has a higher amount of initial assets, a_t, or if the bank relaxes collateral requirements, a_b.

Corollary 1. *Relaxing credit constraints in the form of a reduction in the non-interest costs of borrowing or reducing collateral requirements increase the probability of Type-1 self-employment.*

Second, if the paid-employed worker will be allowed to take a one-period non-business loan (e.g., consumer loan, mortgage), then the Type-2 self-employed worker who initially did not have access to the credit market after a change in the non-interest cost of borrowing may have incentives to switch to paid-employment in order to use credit products. In the Appendix,

[10] While it is more natural to think about the credit market accessibility through the interest costs of borrowing, I am concerned with the non-interest costs rather than the interest costs.

I show that the probability of switching from Type-2 self-employment at period t to paid-employed at period $t+1$ increases if the non-interest cost of borrowing ψ_t decreases.

$$\frac{\partial Pr\left(SE_{2t} \rightarrow PE_{t+1}\right)}{\partial \psi_{t+1}} = \frac{\partial Pr\left(\delta(j) = \{SE_{2t}, PE_{t+1}\}\right)}{\partial \psi_{t+1}} \geq 0$$

Corollary 2. *Relaxing credit constraints in the form of a reduction of non-interest costs of borrowing increases the probability of switching from Type-2 self-employment to paid-employment.*

4 Empirical Model

In this section, I discuss the timeline of the empirical model and derive estimable equations from the theoretical model that allows for the estimation of the causal effect of credit market accessibility on the transition into self-employment.

The Timeline of the Model

In the empirical model at the beginning of period t, the individual i can be in one of four employment states j, Y_{ijt}: paid-employment ($j = 0$), Type-1 self-employment ($j = 1$), Type-2 self-employment ($j = 2$) or non-employment ($j = 3$). The individual i also has a set of time-varying individual characteristics X_{it} (e.g., age, years of schooling, marital status, the value of large, durable assets and etc.) and constant individual characteristics X_t (e.g., gender, race and Hispanic ethnicity). The agent observes local credit market accessibility (the proximity of bank services) at period t, Z_{it}, that is a set of the following variables[11]: the average distance from the individual to the 10 nearest banks, the number of banks and branches withing 5 miles of an individual's home, and the number of bank workers per 1000 pop. at the county level. The proximity of the bank may create incentives for the individual i to change a labor state. For example, if the use a bank becomes cheaper for a paid-employed person, she may make decision to become self-employed and open a new business; or if the self-employed individual i does not satisfy bank requirements (e.g., annual income requirements, collateral requirements, credit score, etc.), she can make decision to switch to paid-employment to be able to use bank services. At the beginning of period $t+1$, an econometrician observes a new employment state for the individual i. The purpose of this empirical analysis is to show how the proximity of

[11] These variables were aggregated into a standardized z-scores index. The z-score index is scaled such that a higher value means better access to the formal credit market (e.g., shorter borrower-lender distance). And, for the convenience of interpretation, I standardize the credit market accessibility index with a mean of zero and a standard deviation of one.

bank Z_{it}, with explanatory variables X_{it} and X_i influence the transition between labor states $Y_{ijt} \rightarrow Y_{ijt+1}$, and I model this transition via the dynamic multinomial logit process described below.

Dynamic Multinomial Logit Model of Employment

The purpose of this empirical analysis is to show how the proximity of bank, Z_{it}, and explanatory variables, X_{it} and X_i, influence the transition between labor states $Y_{ijt} \rightarrow Y_{ijt+1}$. For empirical purposes, I assume that the choice specific error term from the theoretical model can be decomposed as time-invariant and time-variant unobservables, $\epsilon^j_{it+1} = \mu_{ij} + \upsilon_{ijt+1}$, where μ_{ij} captures time-invariant unobservables, and υ_{ijt+1} follows the Type 1 extreme value distribution. Also, the utility function from the theoretical model has a linear parametrization. The probability of choice j is:

$$Pr_{t+1}(j|X_{it}, Z_{it}) = \frac{exp(Y_{ijt}\beta_{1j} + Z_{it}\beta_{2j} + X_{it}\beta_{3j} + \mu_{ij})}{\sum_{j' \in \{PE,SE_1,SE_2,NE\}} exp(Y_{ij't}\beta_{1j'} + Z_{it}\beta_{2j'} + X_{it}\beta_{3j'} + \mu_{ij'})} \tag{1}$$

The probability of employment choice j depends on the previous period job type Y_{ijt}, and the coefficient β_{1j} shows the mobility of individuals across employment states between t and $t+1$. As a proxy for the non-interest cost of borrowing, I use the measure of credit market accessibility Z_{it} (later CMA). A higher value of CMA implies closer proximity to banking services. Since banks fulfill many functions, I assume that future unobserved shocks do not influence the availability of banking services and an individual's job choice in a given year. The measure of bank availability Z_{it} is assumed to be exogenous conditional on the time-constant unobserved effect (μ_{ij}) and time-varying observed factors that may influence the opening of new bank offices in the area, but I discuss later potential solutions to possible violations of this assumption. The coefficient β_{2j} is another coefficient of interest that shows the direct effect of credit market accessibility on the type of employment. The agent ability θ, the amount of available assets a_{t+1}, and the wage for a salary employed worker w_{t+1} are included in X_{it}, and I will discuss actual measures for these variables later.

The model can be re-written in a more conventional log-odds form by choosing the non-employed type as the base category:

$$ln\frac{Pr(Y_{it+1} = m,\ m \in \{PE, SE_1, SE_2\})}{Pr(Y_{it+1} = NE)} = Y_{ijt}\beta_{1j} + Z_{it}\beta_{2j} + X_{it}\beta_{3j} + \mu_{ij} \tag{2}$$

In the eq.2, I allow μ_{ij} to be correlated across labor states.

Identification

Several issues arise in the estimation of eq.2. The first problem is the possible correlation of unobserved heterogeneity, μ_{ij}, with explanatory variables. Following the literature of Murtazashvili and Wooldridge (2016), Papke and Wooldridge (2008), I add the Mundlak-Chamberlain device or the longitudinal average of time-varying explanatory variables X_{it}, which models permanent unobserved heterogeneity as a linear projection of the time-average of time-varying characteristics. This approach accounts for the endogeneity of inputs with regards to unobserved factors like an individual fixed-effect model. Since used panel dataset is unbalanced, it is impossible to use the original approach of Wooldridge (2005) and include a complete history of lagged explanatory variables, $X_t^+ = (X_{i2}, ..., X_{iT})$ and $Z_t^+ = (Z_{i2}, ..., Z_{iT})$. But I can use the modified version of Wooldridge (2005) proposed by Rabe-Hesketh and Skrondal (2013), and authors suggest using the within-means of individual time-varying characteristics based on all periods excluding the first period observations, $\overline{X_i^+} = \frac{1}{T-1}\sum_{t=2}^{T} X_{it}$ and $\overline{Z_i^+} = \frac{1}{T-1}\sum_{t=2}^{T} Z_{it}$. $\overline{X_i^+}$ and $\overline{Z_i^+}$ is analog to the Mundlak-Chamberlain device but without initial values. This approach accounts for the possible endogeneity of CMA with regards to unobserved factors in a manner similar to fixed-effect model.

The second issue is the initial conditions problem that occurs due to the correlation between μ_{ij} and the initial observation Y_{ij1}, and endogenous initial conditions require the specification of conditional distribution for Y_{ij1}. I solve this problem through a solution suggested by Wooldridge (2005) and Rabe-Hesketh and Skrondal (2013). This solution specifies the conditional distribution of μ_{ij} via an auxiliary model that includes the initial dependent variable, Y_{ij1}, the initial-period explanatory variables, X_{i1} and Z_{i1}, and within-means of individual time-varying characteristics based on all periods excluding the first period observations, $\overline{X_i^+}$. The model specifies the following conditional density of unobserved heterogeneity :

$$\mu_{ij} = \pi_{1j}Y_{ij1} + \overline{X_i^+}\pi_{2j} + \pi_{3j}X_{i1} + \overline{Z_i^+}\pi_{4j} + \pi_{5j}Z_{i1} + \eta_{ij} = G_i\pi_j + \eta_{ij} \tag{3}$$

where $\overline{X_i^+}$ is the within-means of individual time-varying characteristics based on all periods excluding the first period observations $\overline{X_i^+} = \frac{1}{T-1}\sum_{t=2}^{T} X_{it}$; Y_{ij1} the initial-period dependent variable; X_{i1} the initial-period explanatory variables; $\overline{Z_i^+}$ is the within-means of the credit market accessibility index based on all periods excluding the first period observations $\overline{Z_i^+} = \frac{1}{T-1}\sum_{t=2}^{T} Z_{it}$; Z_{i1} the initial-period credit market accessibility; η_{ij} is the error term where $Cov(\eta_{ij}; \upsilon_{ijt+1}) = 0$. The vector G_i consists of the initial dependent variable, initial explanatory variables, and within-means of explanatory variables in subsequent periods. The approach suggested by Wooldridge (2005) and Rabe-Hesketh and Skrondal (2013) has several advantages. It does not require instruments, because initial conditions are not modeled separately (in contrast to Heckman (1987) approach). It can be applied to unbalanced panels, and the within-means terms $\overline{X_i^+}$ and $\overline{Z_i^+}$ allow for the correlations between explanatory variables and unobserved heterogeneity μ_{ij}.

The substitution of equation (3) into (2) leads to the standard random-effects multinomial logit model:

$$ln\frac{Pr(Y_{it+1}=m,\ m\in\{PE,SE_1,SE_2)\})}{Pr(Y_{it+1}=NE)}=Y_{ijt}\beta_{1j}+Z_{it}\beta_{2j}+X_{it}\beta_{3j}+G_i\pi_j+\eta_{ij} \qquad (4)$$

To estimate the model with unobserved heterogeneity, η_{ij}, I use two approaches. In the first approach, I assume multivariate normality of the error term, $\eta_{ij}\sim N(0;\sigma_\eta^2)$. In the second approach, I use a variation of the discrete factor approximation (Heckman and Singer, 1984).

Endogenous Locations of Branches. Identification of the causal effect of credit market accessibility on self-employment is challenging because the choice of where to locate a bank is not exogenous. In particular, the decision to open or close a branch is endogenous to local economic conditions that also affect the individual's home location and the decision to become self-employed. There are several solutions to this issue. One possible solution is to control for a large set of variables that proxy local economic conditions at the county level, for example, the number of business establishments, unemployment rate, share of the population with professional, scientific, management and admin. education. For this approach, I add to Eq.4 the set of described variables that define local economic conditions, L_{it}. Also, to avoid the problem of a possible correlation of unobserved individual permanent heterogeneity with local economic conditions variables and initial condition problem, I include analogs with the Mundlak-Chamberlain device, $\overline{L_i^+}=\frac{1}{T-1}\sum_{t=2}^{T}L_{it}$, and the initial period local economic conditions, L_{i1}.

$$ln\frac{Pr(Y_{it+1}=m,\ m\in\{PE,SE_1,SE_2)\})}{Pr(Y_{it+1}=NE)}=Y_{ijt}\beta_{1j}+Z_{it}\beta_{2j}+X_{it}\beta_{3j}+L_{it}\beta_{4j}+F_i\pi_j+\eta_{ij} \qquad (5)$$

where $F_i\pi_j=\pi_{1j}Y_{ij1}+\overline{X_i^+}\pi_{2j}+\pi_{3j}X_{i1}+\overline{Z_i^+}\pi_{4j}+\pi_{5j}Z_{i1}+\overline{L_i^+}\pi_{6j}+\pi_{7j}L_{i1}$.

Equation (5) is the main specification for the empirical part of the paper. I estimate Eq.(5) using both estimation approaches for the unobserved heterogeneity η_{ij}, that I described above.

Second, I estimate Eq.(5) using the lagged measure of credit market accessibility Z_{it-1}. The one-year measure of bank availability Z_{it-1} is assumed to be exogenous conditional on the time-constant unobserved effect (μ_{ij}) and time-varying observed factors that may influence the opening of new bank offices in the area.

The other approach is to estimate model Eq.(5) only for individuals who remained at the same location so any change in the borrower-lender distance was attributable to changes in the bank's location.

The last approach is to use an instrumental variable for changes in the local credit market accessibility. I use Nguyen's (2019) and Garmaise and Moskowitz's (2006) suggested instrumnet which is the exposure to merger-induced consolidation at the Census tract-level[12]. I consider mergers of large banks, defined as mergers in which both banks held at least \$1 billion in pre-merger assets. The average Census tract is 1.5 square meters, and the decision of large banks to merge is plausibly exogenous to local economic conditions in a Census tract where both banks have a branch. The merge-induced consolidation decreases bank competition at the tract level, which may be followed by a branch closure (Nguyen, 2019), an increase in the borrower-lender distance, a decrease in the number of branch employees and a decline in the local credit market accessibility.

The empirical framework for the instrumental variable approach is the dynamic multinomial logit model with the CMA measure, Z_{it}, treated as endogenous:

$$\begin{cases} ln\frac{Pr(Y_{it+1}=m,\ m\in\{PE,SE_1,SE_2)\})}{Pr(Y_{it+1}=NE)} = Y_{ijt}\beta_{1j} + Z_{it}\beta_{2j} + X_{it}\beta_{3j} + L_{it}\beta_{4j} + F_i\pi_j + \eta_i\beta_{5j} \\ Z_{it} = E_{it}\gamma_1 + Y_{ijt}\gamma_2 + X_{it}\gamma_3 + L_{it}\gamma_4 + G_{it} + \eta_i \end{cases} \tag{6}$$

where E_{it} equals 1 if the individual i lives in a Census tract where two large banks merged in year t, $F_i\pi_j = \pi_{1j}Y_{ij1} + \overline{X_i^+}\pi_{2j} + \pi_{3j}X_{i1} + \overline{Z_i^+}\pi_{4j} + \pi_{5j}Z_{i1} + \overline{L_i^+}\pi_{6j} + \pi_{7j}L_{i1}$, $G_{it} = Y_{ij1}\psi_1 + \overline{X_i^+}\psi_2 + X_{i1}\psi_3 + Z_{i1}\psi_4 + L_{i1}\psi_5 + \overline{L_i^+}\psi_6$. For computational simplicity, I assume that $\eta_i \in N(0;1)$.

Eq.(6) jointly estimates the individual's i labor decisions due to changes in CMA, Z_{it}, and CMA as a function of individual characteristics,X_{it}, current labor state, Y_{ijt}, local economic conditions, L_{it}, and instrumental variable E_{it}. The error term, η_i, contains permanent individual unobserved heterogeneity that influences both the labor decisions and credit market accessibility for individual i.

[12] In the literature, it is accepted to use overidentified models to assess the validity of IV. In this paper, I use the exact identified model, and it is impossible to validate instrumentation (Wooldridge (2016)), but for further research, I would like to use more than one instrumental variable. However, some papers stay that the validity of the IV implied by the economic model is an identifying assumption that cannot be tested (Parente and Silva (2012), Deaton (2010)). Because the validity of the overidentifying restrictions is neither sufficient nor necessary for the validity of the IV implied by the underlying economic model, and therefore provide little information on the possibility of identifying the parameters of interest.

5 Data

Sources of Data

This paper combines several sources of US data. The primary source is the Community Advantage Panel Survey (CAPS) database collected by the UNC Center for Community Capital. CAPS comprises 11 years of panel survey data provided by approximately 5,000 low- and moderate–income homeowners and renters during the period of 2003-2014. The homeowners recruited to participate in CAPS received mortgages between 1999 and 2003 through the Community Advantage Program, and the participating renters were recruited to match these homeowners with respect to geographic proximity and income ceilings. The survey collects a wide variety of information, including demographics and family formation, mobility and housing tenure choice, unemployment, wealth and asset holdings, social capital and civic engagement, and housing satisfaction. The CAPS database was used in several works (e.g., Quercia et al. (2011), Manturuk et al. (2017))[13]. The data Appendix provides more information about the survey design.

Data on career choices and self-employment. CAPS contains self-reported information about the primary work activity at the time of the survey. The survey has four types of employment: workers in private businesses, government workers, self-employed workers, and workers in a family business. The first two categories were combined as "paid-employed workers" - workers in private company and government workers. Self-employed and workers in the family business are different labor types due to various liabilities; in this way, I consider only self-employed in the analysis and exclude workers in the family business from the data sample.

Data on bank locations. All information about bank offices is from the FDIC Summary of Deposits (SOD) data for 2003-2014 as of June 30 of the corresponding year. The SOD data include the full street addresses of the bank headquarters and its branches, the total amount of assets, and the latitude and longitude of the bank headquarters and each branch bank since 2008. The maps in Figure 5 show the location of all banks and branches from the FDIC SOD database (red dots) in 2003 and 2014. The FDIC SOD contains information about 87279 bank locations in 2003 and 94521 bank locations in 2014. Therefore, the number of bank offices increased, but the offices are more concentrated. To show the difference in concentration between 2003-2014 Fig.6 also has the hot spot analysis (blue areas). The hot spot analysis shows the z-score for the Getis-Ord Gi* statistic for each location in a dataset. Darker areas tell where bank locations with high values cluster spatially.

[13] The full list of publications can be found here `https://communitycapital.unc.edu/files/2017/10/Paper_22929_extendedabstract_1348_0.pdf`

Bank addresses were transferred to geocodes[14] and linked to home locations of CAPS respondents, such that each respondent was linked to each bank within the same state. Bank locations are used to identify the distance[15] from the location of a respondent's home to banks (the minimum distance; distance to nearest 5, 10, or 15 banks); the number of banks per state, number of banks within 5, 10 or 15 miles of the respondent.

Other sources of data. Other sources of data are the U.S. Census Bureau Statistics (County Business Patterns), the Bureau of Labor Statistics' Local Area Unemployment data (county-level unemployment rates and the size of a county's civilian labor force), FDIC Assets and Liabilities report (total number of employees per branch (full-time equivalent)), county-level demographic information from the American Community Survey (ACS), and the Mergers and Acquisitions Database of Federal Reserve Bank of Chicago[16].

Explanatory Variables

As explanatory variables, I use the set of time-invariant individual characteristics, such as gender (=1 if female), race (white, black, other), and Hispanic ethnicity. Also, there are time-varying variables, including age, age squared, years of schooling, spouse's years of schooling, marital status, household size, number of children under the age of 14 currently living in household, region, calendar year effects, the log of the value of large durable assets (houses, land, other real estate and cars), the difference in log-income for wage and salary workers and self-employed (non-incorporated) at the county level. The latter variable serves as a substitute for individual income, which is likely to be endogenous. Unfortunately, CAPS does not contain detailed information about wages or personal income, only aggregated household income.

As a measure of local economic conditions, I use total population at the zip-code level, the share of the population with professional, scientific, management, and admin. education at the zip-code level, the unemployment rate and the number of business establishments at the county level.

[14] Some banks adresses were dropped because their adresses are incomplete. On average the amount of branches with incomplete information is less than 1% of total amount of branches.

[15] For estimating distance I use SAS geocode procedure, because SAS allows to estimate distances without sending sensitive information to external servers, but downside of using SAS is that it estimates the Euclidean distance between 2 locations. Boscoe et al. (2012) conclude that for nonemergency travel the added precision offered by the substitution of travel distance, travel time, or both for straight-line (Euclidean) distance is largely inconsequential.

[16] https://www.chicagofed.org/banking/financial-institution-reports/merger-data

Constructed Variables

Types of self-employed. The biggest challenge is how to identify the types of self-employed for empirical analysis. Unfortunately, from the used CAPS database, it's impossible to identify self-employed businesses without business loans (or with the restricted access to the credit market). In the empirical model, I use several specifications as a proxy for the two types of self-employed. The main specification is based on types of self-employed businesses - incorporated or unincorporated. Levine and Rubinstein (2017) show that businesses of incorporated and unincorporated self-employment demand different skills and abilities: incorporated self-employed have comparatively strong nonroutine cognitive abilities, while unincorporated self-employed are involved in businesses demanding relatively strong manual skills. Later, Levine and Rubinstein (2018) define self-employed types as entrepreneurs and other self-employed, and for empirical analysis, they use incorporated self-employed as a proxy for entrepreneurial self-employed, and unincorporated self-employed as other self-employed. For the main specification, I adopt the approach of Levine and Rubinstein (2018) and use incorporated self-employed as a proxy for Type-1 self-employed, and unincorporated self-employed as a proxy for Type-2 self-employed.

The second possible definition for types of self-employed is based on the degree to which the business demands entrepreneurial abilities. Unfortunately, the database does not contain the assessment of the respondents' abilities. To evaluate the level of entrepreneurial skills, I use an external database O*NET OnLine[17]. The database contains the list of occupations along with the level of entrepreneurship. The level of entrepreneurship has a range between 1 and 5, where level 1 is equivalent to little or no enterprising activities; level 5 is equal to an extensive level of entrepreneurship. Self-employed workers are referred to Type-1 if the level of entrepreneurial activities is equal to 4 or 5; all other self-employed are Type-2.

The third definition of types of self-employed is based on the aggregate level of the following characteristics: problem-solving, creation of new ideas, and thinking. I use characteristics of occupations from the O*NET OnLine database. I aggregate these characteristics into a standardized z-scores index. The higher value of index means a higher level of problem-solving, creation of new ideas and thinking. Self-employed workers with z-score above 0 are referred to Type-1 self-employed, all others - to Type-2 self-employed.

The other definition of types of self-employed originates from studies about labor informality in developing countries. In these studies, it is common to refer to small-scale self-employment as the informal labor sector. Studies from countries with a large informal sectors find that

[17] National Center for O*NET Development, www.onetonline.org/

new bank openings create incentives for individuals working in the informal sector to switch to jobs with verifiable income to be eligible for the loan and, thus, transit to wage employment (Malkova et al., 2018). The share of self-employed individuals with unverifiable income may fall when the borrower-lender distance decreases. Thus, the effect of CMA on the transition to self-employment may differ across individuals. The literature has not explored labor informality broadly in the U.S., because the biggest challenge is the availability of the data, and most American surveys do not contain questions about informal labor activity. The study of Bracha and Burke (2016) finds that 37 percent of non-retired U.S. adults participated in some informal work in 2015. Due to the lack of other data and methodology about labor informality in the U.S., I use the non-participation in earned income tax credit (EITC) program if the individual is eligible for this program as a proxy for possible tax evasion. Type-1 self-employed are the self-employed who participated in EITC, while Type-2 self-employed are the self-employed who did not engage in EITC even if they are eligible for this program. Not eligible for EITC workers are excluded from the sample for estimation of the model for this definition.

Table 1: Alternative Definitions of Types of Self-Employment

	Type 1	Type 2
Based on survey questions		
1. Type of business	Incorporated	Unincorporated
2. Proxy for tax evasion	Participation in EITC by eligible individuals	Non-participation in EITC by eligible individuals
Based on ILO occupation codes (matched to National Center for O*NET dataset)		
3. Level of entrepreneurial abilities	High entrepreneurial abilities	Low entrepreneurial abilities
4. Level of abilities: problem solving, creation of new ideas and thinking	Z-score index >0	Z-score index ≤ 0

Credit market accessibility. The credit market accessibility is measured as an index based on the combination of the following variables: the average distance from the individual to the 10 nearest banks, number of banks and branches within five miles of the individual's home, number of bank workers per 1000 pop. at the county level. These variables were aggregated into a standardized z-scores index. The z-score index is scaled such that a higher value means better access to the formal credit market (e.g., shorter borrower-lender distance). And, for the convenience of interpretation, I standardize the credit market accessibility index with a mean of zero and a standard deviation of one.

Instrumental Variable. To estimate equation (6), I create an instrumental variable, E_{it}, that equals one if the individual i lives in the Census tract where two large banks undergo a merger in year t, and zero otherwise. To create the instrument, I use data on merger activity from the Mergers and Acquisitions Database (M&A) of the Federal Reserve Bank of Chicago. Each "buyer" and "target" bank have an RSSD ID that allows merging the M&A database to FDIC SOD; thereby, I can identify branches that can be exposed to merger-induced consolidation. I consider mergers of large banks, where both banks held at least $1 billion in pre-merger assets. For all branch addresses in FDIC SOD, I identify the Census tract number using SAS geocode procedure. Using this data, I identify tracts where branches of two large banks may be exposed to merger-induced consolidation.

6 Results

In this section I discuss my empirical findings and the implications for how the increase in the CMA affects self-employment. The first step is to verify that my credit market access (CMA) index is a valid measure credit availability in that individuals are more likely to apply for loans and more likely to be approved as CMA increases. Second, I present the results from estimation of the dynamic multinomial logit model of employment with correlated random effects in using both parametric and non-parametric approaches. For the main specification as given in Eq. (5), self-employment is defined on the basis of the legal structure; Type-1 and Type-2 self-employed are owners of incorporated and unincorporated businesses respectively. Third, I use the estimated coefficients to determine the average marginal effects of CMA on the sizes of the four employment sectors and the transition probabilities among the sectors. Partly as a robustness check, I then present the results from other specifications of the model and finally determine how increases in banking services across geographical areas affect self employment.

Access to Credit Products

In this subsection, I want to discuss how the borrower-lender distance influences access to credit products. Unfortunately, CAPS data does not contain information about business-related lending, and in a perfect situation, the influence of CMA on labor decisions should be estimated through the access to credit products. Using available data, I highlight two main groups of variables characterizing the access to credit: the access to different credit products and indicators of reduction in credit access. Variables describing the access to various products are a set of dummy variables: if the respondent applied for a new credit card in the last 12 months, if the respondent applied for a car loan in the previous 12 months, if the respondent opened a new charge card in the last 12 months, if the respondent refinanced a mortgage during the previous 12 months. Indicators of the reduction in credit access are a set of dummy variables: if the respondent's application for a new credit has been denied in the last 12 months, if the respondent's available limit for credit cards was reduced in the last 12 months, if the respondent was asked to pay off the remaining balance for a loan over the previous 12 months, if the respondent filed bankruptcy in the last 12 months.

To evaluate the effect of credit market accessibility on the access to credit, I estimate a random-effects probit model for each of the variables characterizing the access to credit. To avoid a problem of possible correlation of unobserved individual permanent heterogeneity with explanatory variables I include analogs with the Mundlak-Chamberlain device, $\overline{X_i^+} = \frac{1}{T-1}\sum_{t=2}^{T} X_{it}$ and $\overline{L_i^+} = \frac{1}{T-1}\sum_{t=2}^{T} L_{it}$.

$$Pr(C_{it} = 1) = \Phi(Z_{it-1}\beta_1 + X_{it}\beta_2 + L_{it}\beta_3 + \overline{X_i^+}\beta_4 + \overline{L_i^+}\beta_5 + \nu_i) \tag{7}$$

where ν_i are i.i.d., $\nu_i \in N(0, \sigma_\nu^2)$, and Φ is the standard normal cumulative distribution function. C_{it} is a dummy variable that equals 1 if the respondent applied for a new credit card in the last 12 months; if the respondent applied for a car loan in the last 12 months; if the respondent opened a new charge card in the last 12 months; if the respondent refinanced a mortgage in the last 12 months; if the respondent's application for new credit has been denied in the last 12 months; if the respondent's available limit for credit cards was reduced in the last 12 months; if respondent was asked to pay off remaining balance for loan in the last 12 months; if the respondent filed bankruptcy in the last 12 months.

The results of equation (7) are shown in Table 4. Results suggest that improvements in the CMA (smaller borrower-lender distance) increases the risk of receiving a new credit card, having a car loan, obtaining a charge card or refinancing a mortgage. These results support the evidence from the literature that shorter borrower-lender distance makes loans more affordable. The other part of the results suggests that improvements in the CMA decrease the risk of the reduction in access to credit. Shorter borrower-lender distance reduces the risk of denying the application for new credit, reduction in the available limit for credit cards, being asked to pay off the remaining balance, and filing for bankruptcy.

Results of the Main Model

The results of the main model are reported in Table 5. The table presents the relative risk ratios from the dynamic multinomial logit model of employment with correlated random effects.

The dependent variable is the employment status in $t+1$, with the non-employed state chosen as the base outcome and the paid-employment state selected as the omitted lagged dependent category. The table presents the results of the model for the main specification, and the type of self-employment is defined based on the legal status of the business - Type-1 self-employed are owners of an incorporated business, and Type-2 self-employed are owners of an unincorporated business. Table 5 shows both of the estimation approaches for permanent unobserved heterogeneity. In this subsection, I discuss results for the parametric estimation, $\mu_{ij} \in N(0, \sigma_\nu^2)$; the results for the discrete factor approximation are very similar.

Previous labor status plays an essential role in the type of current job. Predictably, for self-employed and non-employed individuals, the risk of working in the paid-employed sector at period $t+1$ relative to paid employed workers is small and negative (the odds are 0.191, 0.181, and 0.040, respectively). For both types of self-employed, the previous job in the self-employed sector increases the risk of being either type of self-employed. The relative risk ratio of being Type-1 self-employed at period $t+1$ relative to paid employed workers at period t for Type-1 self-employed is 50.034, and 13.791 for Type-2 self-employed. Non-employed are less likely to switch to any type of self-employment. The relative risk ratio of being Type-2 self-employed at period $t+1$ relative to paid employed workers at period t for Type-1 self-employed is 12.107, and 23.329 for Type-2 self-employed.

The result of the main interest is the coefficient of the credit market accessibility index. A one-standard-deviation increase in the credit market accessibility index increases the odds of being paid employed by 1.182. At the same time, a one-standard-deviation improvement in CMA increases the relative risk of becoming Type-1 self-employed by 1.323. But the relative risk ratio for Type-2 self-employed is insignificant.

The amount of large, durable assets plays an essential role in being self-employed and increases the risk of being Type-1 self-employed or Type-2 self-employed, but it does not significantly influence the risk of being a paid-employed worker. Also, the difference in log-income for wage and salary workers and self-employed increases the risk of being paid-employed, decreases the risk of being Type-2 self-employed, and does not influence the risk of being Type-1 self-employed.

To show the average marginal effects of the one-unit improvement in the credit market accessibility index on the size of sectors and the transition probabilities, I re-estimate the main model including the interaction terms between the labor state Y_{ijt} and the credit market accessibility index Z_{it}:

$$ln\frac{Pr(Y_{it+1}=m,\ m\in\{PE,SE_1,SE_2)\})}{Pr(Y_{it+1}=NE)}=Y_{ijt}\beta_{1j}+Z_{it}\beta_{2j}+Y_{ijt}*Z_{it}\beta_{3j}+X_{it}\beta_{4j}+ \\ +L_{it}\beta_{5j}+F_i\pi_j+\eta_{ij} \tag{8}$$

where $F_i\pi_j=\pi_{1j}Y_{ij1}+\overline{X_i^+}\pi_{2j}+\pi_{3j}X_{i1}+\overline{Z_i^+}\pi_{4j}+\pi_{5j}Z_{i1}+\overline{L_i^+}\pi_{6j}+\pi_{7j}L_{i1}$.

Table 6 shows the average marginal effect of a one-unit improvement in the credit market accessibility index on the size of sectors. This one-unit improvement in the credit market accessibility index increases the population share of the Type-1 self-employed by 0.22 ppt and reduces the Type-2 self-employed by 0.20 ppt. Table 6 also shows the average marginal effect (AME) of Z_{it} on the probability of switching between employment states. For Type-1 self-employed, the increase in Z_{it} by one standard deviation increases the likelihood of turning from paid-employment to self-employment by 0.08 percentage points. The effect of Z_{it} on the likelihood of staying in the Type-1 self-employment sector is substantial (3.7 ppt improvement). Similarly, there is a substantial positive effect on the probability of moving from Type-2 self-employed (1.8 ppt). For non-employed, the effect is 0.3 ppt. For Type-2 occupations, the effect is the opposite. The improved credit market accessibility (by one standard deviation) decreases the chances of paid-employed workers to become Type-2 self-employed by 0.07 ppt. The effect of Type-1 self-employed is -1.5 ppt. The likelihood of staying in the self-employed sector for Type-2 occupations goes down by 5.1 ppt per unit increase in Z_{it}. The average marginal effects for non-employed individuals are smaller in magnitude, but they are still substantial. The transition probability from non-employed status to the Type-2 self-employed goes down by 1.2 ppt. The transition probabilities to paid-employment are important results. A one standard deviation improvement in CMA decreases the probability of switching from Type-1 self-employment to paid-employment by 1.6 percentage points. But for Type-2 self-employment, the improvement in CMA increases the probability of switching to paid-employment by 3.8 ppt, which supports the evidence from the theoretical model (Corollary 1) that the reduction of non-interest costs of borrowing is likely to increase the probability of switching from Type-2 self-employment to paid-employment.

In addition to average marginal effects (which are calculated across all individuals in the sample), I also estimate the marginal effects at the mean values of covariates and at different values of Z_{it}. These results are plotted in Figure 9. The figure shows the predicted probabilities of being Type-1 or Type-2 self-employed at different values of the credit market accessibility index. Results show a rise in the size of Type-1 self-employed and a decline in Type-2 self-employed, as credit market accessibility improves.

Robustness Analysis

In this section, I discuss the results of several robustness analyses. First, I compare the results of the main model (Eq.5) for different distributional assumptions for unobserved heterogeneity. I estimate the main model (Eq.5) under assumptions about the multivariate normality of the error term and the discrete factor approximation of the error term. As I mentioned in the previous section, results for both distributional assumptions produce similar effects (Table 5). Second, I compare the results of the main model (Eq.5) for different definitions of self-employed types (Table 7). Third, I show results for different identification assumptions for possible endogeneity of CMA, that I described in the "Identification" section (Table 8). Last, I show the results of the main model (Eq.5) for the definition of self-employed types based on possible tax evasion (last three columns in Table 7) and draw comparisons with the literature on developing countries.

Different definitions of self-employed types. The results of the main model (Eq.5) for other definitions of self-employed groups are presented in Table 7. Columns (1)-(3) show the results of the main specification discussed above. Columns (4)-(6) present the results of the model when the definition of types of self-employed is based on the degree to which the business demands entrepreneurial abilities. Columns (7)-(9) present the results of the model when the definition of types of self-employed is based on aggregate characteristics of occupations. Columns (10)-(12) present the results of the model when the definition of types of self-employed is based on participation in the EITC program (I discuss these results later). In all specifications, a one-standard-deviation improvement in CMA increases the relative risk of becoming Type-1 self-employed and paid-employed.

Solutions to possible endogeneity of CMA. Tables 8 and 9 show results for different identification assumptions discussed in the "Identification" section. The result of main interest is the coefficient of the credit market accessibility index. For the base specification under an assumption about exogenous CMA, a one-standard-deviation increase in the credit market accessibility index increases the odds of being paid-employed by 1.103. At the same time, a one-standard-deviation improvement in CMA increases the relative risk of becoming Type-1 self-employed by 1.197. But the relative risk ratio for Type-2 self-employed is insignificant. Results for other identification assumptions are similar: including lag of CMA, estimating the model only for individuals who did not move including LEC, and including local economic condition (withing-means of LEC and initial conditions). It is interesting that the results for the last specification, including the Mundlak-Chamberlain device for CMA and the initial value of CMA, show even higher relative risk ratios. A one-standard-deviation increase in the credit market accessibility index increases the odds of being paid-employed by 1.207, and a one-standard-deviation improvement in CMA increases the relative risk of becoming Type-1 self-employed by 1.376.

Results for eq.(6), which jointly estimates the individual's i labor decisions due to changes in the CMA, and CMA as a function of local economic conditions and instrumental variable, are shown in Table 9. The instrumental variable equals one if the individual i lives in the Census tract where two large banks undergo a merger in year t. Results show that the exposure to merger-induced consolidation negatively influences credit market accessibility. A one-standard-deviation increase in the credit market accessibility index increases the odds of being paid-employed by 1.097. At the same time, a one-standard-deviation improvement in CMA increases the relative risk of becoming Type-1 self-employed by 1.188. But, the relative risk ratio for Type-2 self-employed is insignificant.

Tax Evasion and Labor Informality. The last three columns of Table 7 provide the results for estimation of the main model (Eq.5) with the definition of types of self-employed based on participation in the EITC program. I use non-participation in the earned income tax credit (EITC) program if the individual is eligible for this program as a proxy for possible tax evasion[18]. Results show that a one-standard-deviation improvement in CMA increases the relative risk of becoming Type-1 self-employed and paid-employed, but decreases the risk of being Type-2 self-employed. These results support the evidence from studies about labor informality in developing countries (Malkova et al., 2019). Unfortunately, the literature has not explored labor informality broadly in the U.S. Bracha and Burke (2016) find that in 2015 37 % of non-retired U.S. adults participated in some type of informal work. The biggest challenge is the availability of the data because primary American surveys don't contain questions about informal labor activity. Bracha et al. (2015)[19] mention that the SCE-SIWP (the Survey of Informal Work Participation) is the only survey of informal work participation that covers a nationally representative sample.

[18]Participants in EITC program have a higher chance of a tax audit.

[19]Authors find that informal work does embody a significant amount of labor market slack. In particular, they find that formal labor market conditions are negatively correlated with informal work hours, and those who are employed part-time for economic reasons are especially likely to turn to informal work, as might be expected if informal hours represent a substitute for formal work.

Policy Simulations

To examine the marginal effect of each component of the credit market accessibility index, I replaced the aggregate index with three measures (the average distance to the 10 nearest banks, the number of banks within 5 miles of a respondent's home, the number of bank workers per 1000 population) and then re-estimate the main model (eq.5). Results are reported in Table 10. Some of the marginal effects for these measures are not significant due to possible multicollinearity issues (the average distance to 10 nearest banks and the number of banks within 5 miles of a respondent's home are correlated).

I report the results of four simulated policies in Table 10. Table 10 shows the predicted probabilities of being paid-employed, Type-1 self-employed, and Type-2 self-employed workers before and after a given policy is implemented. The predicted probability of being non-employed is $1 - \widehat{Pr_{PE}} - \widehat{Pr_{SE_1}} - \widehat{Pr_{SE_2}}$, but it is not shown. Predicted probabilities are evaluated at the starting values indicated in the table and at sample means of all other covariates.

The first policy is to increase banking access for those who have no banking offices within five miles of their homes[20], and specifically to increase the number of branches to 37, because the average number of banks within 5 miles distance from individual home is 36.7 in my sample. Based on the model estimates, I find a statistically significant increase of 0.4 ppt in the number of Type-1 self-employed by 0.4 ppt and a statistically significant decrease of 0.03 ppt in the number of Type-2 self-employed.

The second policy is the use of online banking. To estimate this policy, I look at the changes in the average distance to the 10 nearest banks from the mean value to zero. The effects of this policy are more extensive in magnitude. The estimates predict a substantial increase in the share of the Type-1 self-employed sector by 0.9 ppt, and a decline in the share of the Type-2 self-employed sector by 3.1 ppt. The paid-employment sector is growing by almost 9 ppt, which can be explained by the switching of Type-2 self-employed and non-employed to paid-employment. Usage of online banking is correlated with higher credit market accessibility, and higher credit market accessibility can be a signal of better local economic conditions. Fig. 7 shows the number of online banking users by county from the US Census and locations of all banks in North Carolina. It's clear from the figure that a higher number of online banking users is associated with higher credit market accessibility, which explains the increase in the share of paid-employed workers.

The last policy is an increase in the number of bank workers from 2 per 10,000 pop. (the

[20]The percent of the individuals with no banks withing five miles is 3.1% in the sample.

average number of bank workers) to 5 per 10,000 pop. I also find a substantial effect on employment composition – an increase in the share of the Type-1 self-employed sector by 0.04 ppt and a decrease in the share of the Type-2 self-employed sector by 0.11 ppt.

Overall, policy simulations show strong support for the reduction in Type-2 self-employment and the increase in Type-1 self-employment in response to better credit market accessibility.

7 Conclusion

The paper investigates the role of the credit market channel in the decrease of self-employment in the US. To analyze the impact of credit market accessibility on entry into self-employment, I develop a three-sector Roy model that differentiates between two types of self-employed: entrepreneurs who have growth-oriented businesses that demand physical capital and business loans (Type-1) and other self-employed (Type-2). The focus of the study is the physical accessibility of banks, and I use a novel data source (the Community Advantage Panel Survey database) that allows me to measure the proximity of credit market institutions for all respondents in the data. The combination of the following variables creates the index of physical availability of banks: the average distance to the 10 nearest banks of the location of respondent's home, number of banks within 5 miles of the location of a respondent's home, and the average number of workers per bank's office at the county level.

The empirical study estimates how the selection into self-employment can be driven by credit market institutions using the dynamic multinomial logit model of employment with correlated random effects. The analysis shows the heterogeneous response to changes in credit market accessibility (CMA) among two types of self-employed. Type-1 self-employed workers are owners of growth-oriented businesses that demand physical capital and business loans, and the expansion in CMA increased transition to self-employment. The one-unit improvement in credit market accessibility index increases the population share of the Type-1 self-employed by 0.22 ppt.

Type-2 self-employed are owners of less capital-intensive businesses, and the increase in CMA decreased the transition to this self-employment type. The one-unit improvement in the credit market accessibility index decreases the population share of the Type-2 self-employed by 0.20 ppt. These results can be interpreted oppositely, that the recent decline in physical credit market accessibility shrinks the size of the Type-1 self-employed sector and increases the size of the Type-2 self-employed sector, and the cumulative marginal effect is negative, which indicates that overall self-employment is decreasing. Type-2 self-employed workers switch to paid-employment to be able to use credit products when banks increase their presence near the agent's place of living. This result supports the evidence from studies of countries with a large informal sector which show that new bank openings create incentives for individuals working in the informal sector to switch to jobs with verifiable income to be eligible for a loan and, thus, transfer to salary employment (Malkova et al., 2019). Labor informality has not been broadly explored in the U.S., and this evidence can be widely discovered in further research.

References

Adelino, M., Schoar, A., and Severino, F. (2015). House prices, collateral, and self-employment. *Journal of Financial Economics*, 117(2):288–306.

Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. *The Review of Financial Studies*, 23(7):2757–2788.

Ai, C. and Chen, X. (2003). Efficient estimation of models with conditional moment restrictions containing unknown functions. *Econometrica*, 71(6):1795–1843.

Bajari, P., Chernozhukov, V., Hong, H., and Nekipelov, D. (2015). Identification and efficient semiparametric estimation of a dynamic discrete game. Technical report, National Bureau of Economic Research.

Bates, T. and Robb, A. (2013). Greater access to capital is needed to unleash the local economic development potential of minority-owned businesses. *Economic Development Quarterly*, 27(3):250–259.

Beckhusen, J. (2014). Employment transitions among the self-employed during the great recession. Technical report.

Berger, A. N. and Udell, G. F. (1998). The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of banking & finance*, 22(6-8):613–673.

Bernhardt, A. (2014). Labor standards and the reorganization of work: Gaps in data and research.

Block, J. and Sandner, P. (2009). Necessity and opportunity entrepreneurs and their duration in self-employment: evidence from german micro data. *Journal of Industry, Competition and Trade*, 9(2):117–137.

Boscoe, F. P., Henry, K. A., and Zdeb, M. S. (2012). A nationwide comparison of driving distance versus straight-line distance to hospitals. *The Professional Geographer*, 64(2):188–196.

Bracha, A. and Burke, M. A. (2016). Who counts as employed?: informal work, employment status, and labor market slack.

Bracha, A., Burke, M. A., and Khachiyan, A. (2015). Changing patterns in informal work participation in the united states 2013-2015. *Federal Reserve Bank of Boston Research Paper Series Current Policy Perspectives Paper*, (15-10).

Brevoort, K. P., Wolken, J., and Holmes, J. (2010). Distance still matters: the information revolution in small business lending and the persistent role of location, 1993-2003.

Buera, F. J. (2009). A dynamic model of entrepreneurship with borrowing constraints: theory and evidence. *Annals of finance*, 5(3-4):443–464.

Cagetti, M. and De Nardi, M. (2006). Entrepreneurship, frictions, and wealth. *Journal of political Economy*, 114(5):835–870.

Corradin, S. and Popov, A. (2015). House prices, home equity borrowing, and entrepreneurship. *The Review of Financial Studies*, 28(8):2399–2428.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of economic literature*, 48(2):424–55.

Decker, R., Haltiwanger, J., Jarmin, R., and Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. *Journal of Economic Perspectives*, 28(3):3–24.

Degryse, H. and Ongena, S. (2005). Distance, lending relationships, and competition. *The Journal of Finance*, 60(1):231–266.

Dell'Ariccia, G. and Marquez, R. (2004). Information and bank credit allocation. *Journal of Financial Economics*, 72(1):185–214.

Demyanyk, Y. (2008). Us banking deregulation and self-employment: A differential impact on those in need. *Journal of Economics and Business*, 60(1-2):165–178.

DeYoung, R., Glennon, D., and Nigro, P. (2008). Borrower–lender distance, credit scoring, and loan performance: Evidence from informational-opaque small business borrowers. *Journal of Financial Intermediation*, 17(1):113–143.

Evans, D. S. and Jovanovic, B. (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of political economy*, 97(4):808–827.

Fairlie, R. W. and Krashinsky, H. A. (2012). Liquidity constraints, household wealth, and entrepreneurship revisited. *Review of Income and Wealth*, 58(2):279–306.

Fan, W. and White, M. J. (2003). Personal bankruptcy and the level of entrepreneurial activity. *The Journal of Law and Economics*, 46(2):543–567.

Fort, T. C., Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). How firms respond to business cycles: The role of firm age and firm size. *IMF Economic Review*, 61(3):520–559.

Fracassi, C., Garmaise, M., Kogan, S., and Natividad, G. (2013). How much does credit matter for entrepreneurial success in the united states? *Available at SSRN.*

Freixas, X. and Rochet, J.-C. (2008). *Microeconomics of banking.* MIT press.

Garmaise, M. J. and Moskowitz, T. J. (2006). Bank mergers and crime: The real and social effects of credit market competition. *the Journal of Finance*, 61(2):495–538.

Gentry, W. M. and Hubbard, R. G. (2004). Entrepreneurship and household saving. *Advances in economic analysis & policy*, 4(1).

Greenstone, M., Mas, A., and Nguyen, H.-L. (2014). Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and 'normal'economic times. Technical report, National Bureau of Economic Research.

Guiso, L., Sapienza, P., and Zingales, L. (2004). Does local financial development matter? *The Quarterly Journal of Economics*, 119(3):929–969.

Hauswald, R. and Marquez, R. (2006). Competition and strategic information acquisition in credit markets. *The Review of Financial Studies*, 19(3):967–1000.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, pages 271–320.

Heckman, J. J. (1987). *The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process and some Monte Carlo evidence*. University of Chicago Center for Mathematical studies in Business and Economics.

Herkenhoff, K., Phillips, G., and Cohen-Cole, E. (2016). The impact of consumer credit access on employment, earnings and entrepreneurship. Technical report, National Bureau of Economic Research.

Hipple, S. F. (2016). Self-employment in the united states. *Monthly Labor Review*, 133(9):17–32.

Hollander, S. and Verriest, A. (2016). Bridging the gap: the design of bank loan contracts and distance. *Journal of Financial Economics*, 119(2):399–419.

Holtz-Eakin, D., Joulfaian, D., and Rosen, H. S. (1994). Sticking it out: Entrepreneurial survival and liquidity constraints. *Journal of Political economy*, 102(1):53–75.

Hurst, E. and Lusardi, A. (2004). Liquidity constraints, household wealth, and entrepreneurship. *Journal of political Economy*, 112(2):319–347.

Johnson, S. (2016). Are mortgage regulations affecting entrepreneurship?

Kashian, R. D., Tao, R., and Drago, R. (2018). Bank deserts in the usa and the great recession: geography and demographics. *Journal of Economic Studies*, 45(4):691–709.

Keeton, W. R. (2017). *Equilibrium credit rationing*. Routledge.

Kim, P. H., Aldrich, H. E., and Keister, L. A. (2006). Access (not) denied: The impact of financial, human, and cultural capital on entrepreneurial entryin the united states. *Small business economics*, 27(1):5–22.

Knyazeva, A. and Knyazeva, D. (2012). Does being your bank's neighbor matter? *Journal of Banking & Finance*, 36(4):1194–1209.

Lederer, P. J. and Hurter Jr, A. P. (1986). Competition of firms: discriminatory pricing and location. *Econometrica: Journal of the Econometric Society*, pages 623–640.

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

Levine, R. and Rubinstein, Y. (2018). Selection into entrepreneurship and self-employment. Technical report, National Bureau of Economic Research.

Lipman, B. J., Wiersch, A. M., et al. (2018). Browsing to borrow:"mom & pop" small business perspectives on online lenders. *Special Reports*.

Malkova, A., Peter, K., and Svejnar, J. (2018). Informality and credit market development.

Manturuk, K. R., Lindblad, M. R., and Quercia, R. G. (2017). *A place called home: The social dimensions of homeownership*. Oxford University Press.

Murtazashvili, I. and Wooldridge, J. M. (2016). A control function approach to estimating switching regression models with endogenous explanatory variables and endogenous switching. *Journal of Econometrics*, 190(2):252–266.

Nguyen, H.-L. Q. (2019). Are credit markets still local? evidence from bank branch closings. *American Economic Journal: Applied Economics*, 11(1):1–32.

Papke, L. E. and Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, 145(1-2):121–133.

Parente, P. M. and Silva, J. S. (2012). A cautionary note on tests of overidentifying restrictions. *Economics Letters*, 115(2):314–317.

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

Partners, M. (2017). State of independence in america.

Petersen, M. A. and Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *The journal of finance*, 49(1):3–37.

Petersen, M. A. and Rajan, R. G. (2002). Does distance still matter? the information revolution in small business lending. *The journal of Finance*, 57(6):2533–2570.

Qian, J. and Strahan, P. E. (2007). How laws and institutions shape financial contracts: The case of bank loans. *The Journal of Finance*, 62(6):2803–2834.

Quercia, R. G., Freeman, A., and Ratcliffe, J. (2011). *Regaining the dream: How to renew the promise of homeownership for America's working families*. Brookings Institution Press.

Rabe-Hesketh, S. and Skrondal, A. (2013). Avoiding biased versions of wooldridge's simple solution to the initial conditions problem. *Economics Letters*, 120(2):346–349.

Rupasingha, A. and Wang, K. (2017). Access to capital and small business growth: evidence from cra loans data. *The Annals of Regional Science*, 59(1):15–41.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, pages 141–156.

Schweitzer, M. E. and Barkley, B. (2017). Is' fintech'good for small business borrowers? impacts on firm growth and customer satisfaction.

Schweitzer, M. E. and Shane, S. (2016). The ins and outs of self-employment: An estimate of business cycle and trend effects.

Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *The Journal of Finance*, 62(2):629–668.

Wiersch, A. M. and Shane, S. (2013). Why small business lending isn't what it used to be. *Economic Commentary*, (2013-10).

Wilmoth, D. and DePriest, D. L. (2016). The missing millennial entrepreneurs. *US Small Business*.

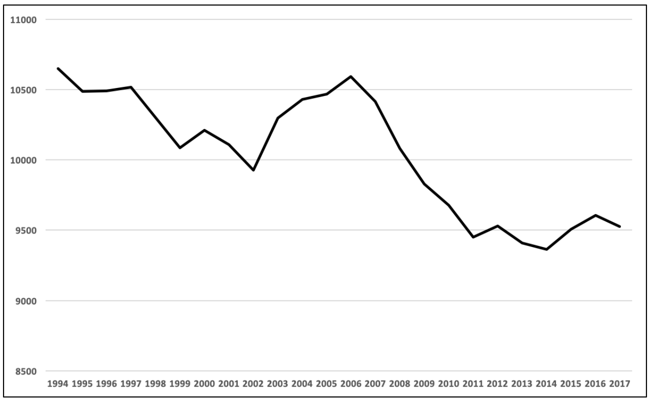
Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics*, 20(1):39–54.

Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach*. Nelson Education.

8 Figures

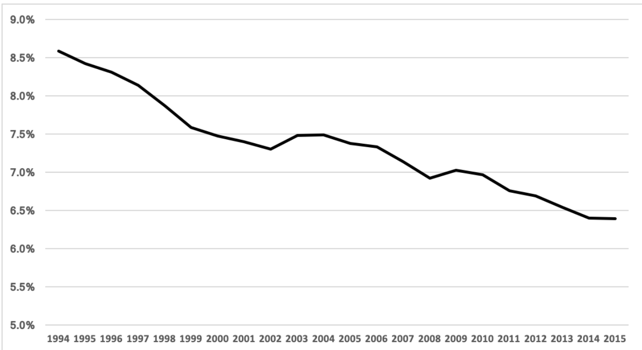
Self-employment in the U.S.

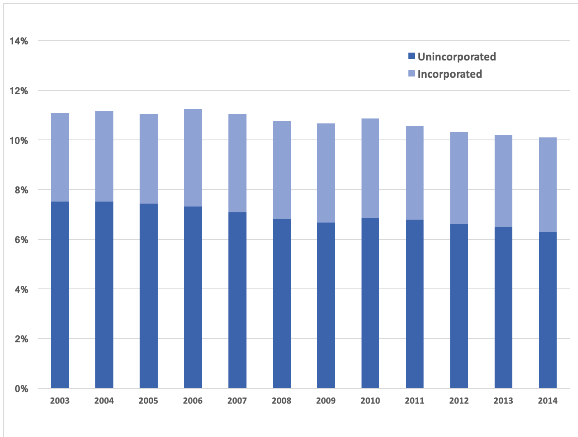
Figure 1: Number of the Unincorporated Self-Employed in the U.S.



Notes. Figure shows number of self-employed from the Current Population Survey (thousands, seasonally adjusted).

Figure 2: The Self-employment Rate in the U.S.

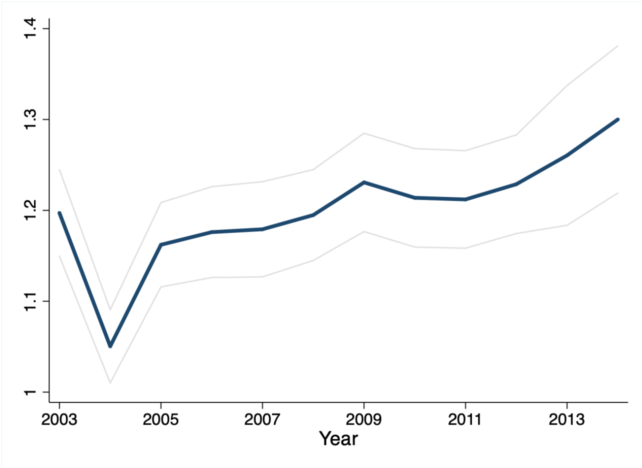




Notes. Left figure shows the self-employment rate as the proportion of total employment from OECD (2019), Self-employment rate (indicator). doi: 10.1787/fb58715e-en (seasonally adjusted). Right figure shows the self-employment rate as the proportion of total employment made up of unincorporated and incorporated self-employed workers from the Current Population Survey (thousands, seasonally adjusted).

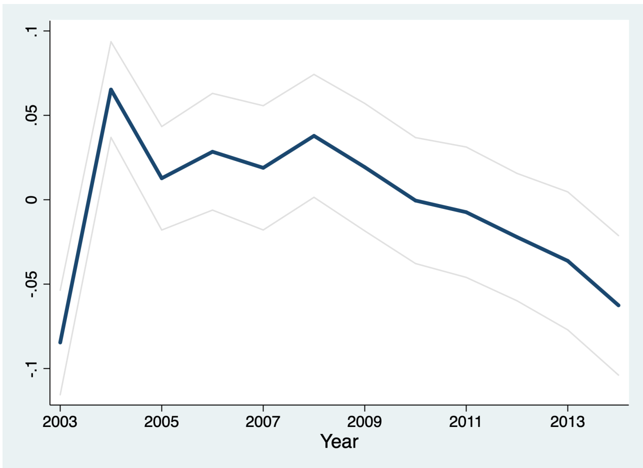
The Credit Market Accessibility in the U.S.

Figure 3: The Average Minimum Distance Between a Borrower and a Lender



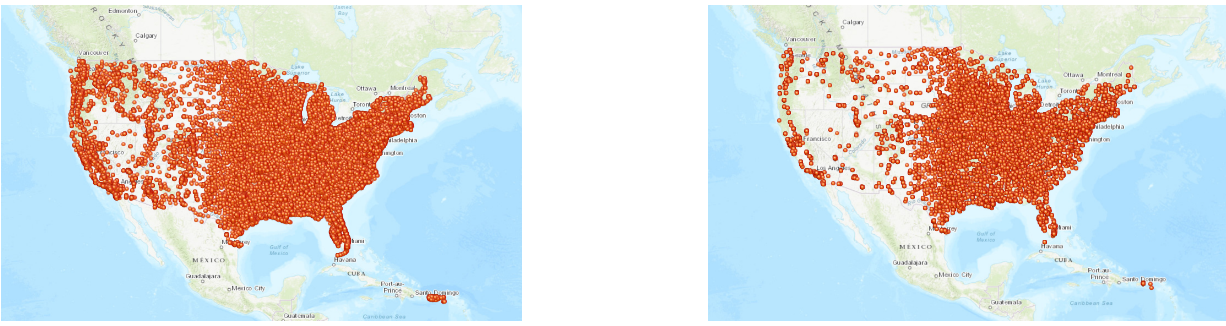
Notes. Figure shows the average minimum distance between a borrower and a lender estimated from the CAPS and the FDIC SOD (miles).

Figure 4: The Credit Market Accessibility Index



Notes. Figure shows the credit market accessibility index. The credit market accessibility is measured as an index based on the combination of following variables: average distance from the individual to 10 nearest banks, average number of the banks and branches withing 5 miles from individual's home, the number of bank workers per 1000 pop. at county level.

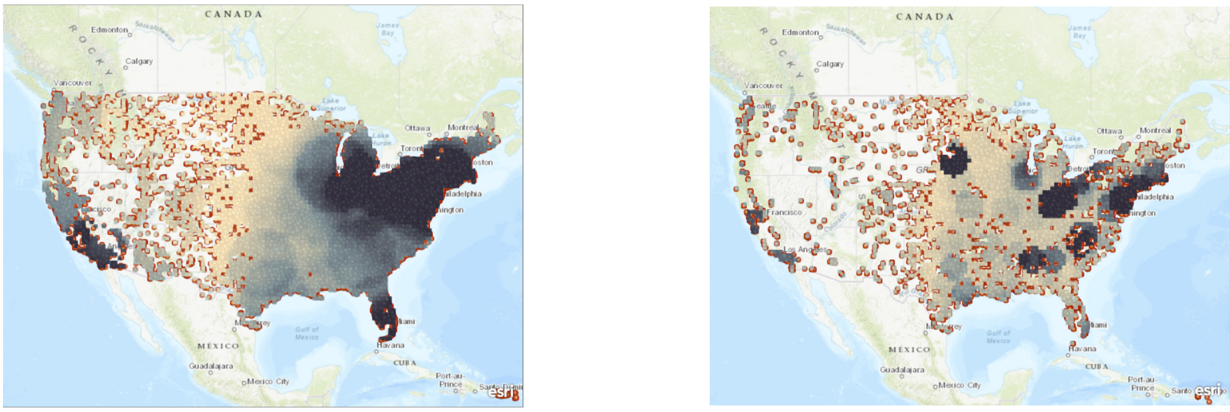
Figure 5: The Location of All Banks in the U.S.



a) 2003 b) 2014

Notes. Figures show the location of all banks and branches form the FDIC SOD database (red dots). In 2003 there are 87279 bank locations in the FDIC SOD, in 2014 there are 94521 bank locations.

Figure 6: The Location of All Banks in the U.S. and the Hot Spot Analysis



a) 2003 b) 2014

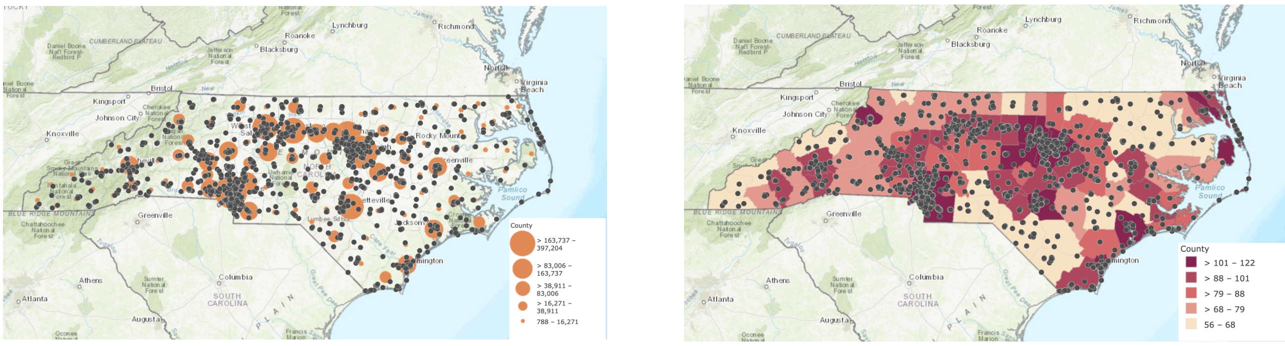
Notes. Figures show the location of all banks and branches form the FDIC SOD database (red dots) and the hot spot analysis made in the ArcGis (blue areas). In 2003 there are 87279 bank locations in the FDIC SOD, in 2014 there are 94521 bank locations. The hot spot analysis shows z-score for the Getis-Ord Gi* statistic for each location in a dataset. Darker areas tell where bank locations with high values cluster spatially.

The z-score for the Getis-Ord Gi* statistic was calculated using ArcGIS tools. The formula for z-score is:

$$G_i^* = \frac{\sum_{j=1}^n w_{i,j}x_j - \frac{\sum_{j=1}^n x_j}{n}\sum_{j=1}^n w_{i,j}}{\sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - \frac{\sum_{j=1}^n x_j}{n}}\sqrt{\frac{n\sum_{j=1}^n w_{i,j}^2 - \left(\sum_{j=1}^n w_{i,j}\right)^2}{n-1}}}$$

where x_j is the attribute value of feature j, $w_{i,j}$ is the spatial weight between features i and j, n is equal to total numbers of features. To be a statistically significant hot spot, a feature will have a high value and be surrounded by other features with high values as well.

Figure 7: Online banking and location of banks in NC, 2018



a) b)

Notes. Figures show the location of all banks and branches in North Carolina form the FDIC SOD database (grey dots) in 2017 and the market potential that an adult has used online banking in the past year in the U.S. in 2018. Figure (a) shows number of adults used online banking in county (biger circle means higher number of users). Figure (b) shows the index created based on the market potential that an adult has used online banking in the past year (darker area means higher market potential).

9 Tables

Table 2: Summary statistics

	Paid -E	Type-1 SE	Type-2 SE	NE
Age**	39.51	40.02	41.11	42.68
	(10.28)	(9.66)	(9.82)	(11.61)
Female***	0.53	0.37	0.48	0.66
Married***	0.58	0.74	0.63	0.45
Number of kids***	1.03	1.34	1.05	1.08
	(1.18)	(1.34)	(1.25)	(1.25)
Number of HH members*	1.88	1.97	1.91	1.91
	(0.78)	(0.67)	(0.81)	(0.95)
Race: White***	0.64	0.76	0.68	0.50
Black	0.24	0.13	0.16	0.33
Other***	0.12	0.11	0.16	0.17
Ethnicity: Hispanic*	0.14	0.14	0.17	0.17
Years of schooling	14.73	15.01	14.71	14.27
	(3.52)	(3.56)	(3.52)	(4.08)
Distance to the nearest bank**, miles	1.22	1.47	1.29	1.1
	(1.56)	(1.78)	(1.62)	(1.39)
Aver. distance to 10 nearest banks***, miles	2.84	3.29	2.89	2.48
	(2.85)	(2.79)	(2.62)	(2.49)
Number of banks within 5 miles***	36.13	40.43	32.21	31.72
	(40.04)	(59.30)	(42.27)	(42.24)
Number of bank workers per 1000 pop.	0.29	0.32	0.30	0.26
	(0.14)	(0.26)	(0.12)	(0.22)
Total assets***, 2014 $	15 801.6	43 507.91	25 598.05	9 019.31
	(45 536.39)	(118 841.8)	(73 625.28)	(35 357.12)
CMA*** (z-score)	0.03	0.018	0.017	-0.17
	(0.97)	(1.01)	(1.06)	(1.01)

Notes. This table shows the summary statistics. Stars (*) shows the p-value for the t-test of mean differences between Type-1 and Type-2 self-employed, where *** denotes significance at 1% level, ** means significance at 5% level, * means significance at 10% level. Number of banks per 1000 population is calculated at the state level. The table indicates that the distance are in miles. Total assets is the sum of non-housing and housing assets adjusted for inflation.

Table 3: Transition matrix between labor markets. Definition based on the type of business: incorporated/unincorporated

	Paid-employed at *t+1*	Type-1 self-employed at *t+1*	Type-2 self-employed at *t+1*	Non-employed at *t+1*	Total at *t+1*
Paid-employed at t	*94.38%*	*0.72%*	*1.08%*	*3.82%*	*84.7%*
	19467	148	223	788	20626
Type-1 self-employed at t	*21.68%*	*56.26%*	*17.76%*	*4.3%*	*2.20%*
	116	301	95	23	535
Type-2 self-employed at t	*23.09%*	*16.59%*	*55.12%*	*5.19%*	*2.85%*
	160	115	382	36	693
Non-employed at t	*26.7%*	*1.08%*	*1.32%*	*70.9%*	*10.26%*
	667	27	33	1771	2498
Total at t	*83.81%*	*2.43%*	*3.01%*	*10.75%*	*100%*
	20410	591	733	2618	24352

Notes. This table shows the average annual probabilities of transitioning from status j at time t to status k in t+1. The definition of types of self-employed is based on the type of business. Type-1 self-employed are owners of incorporated businesses, Type-2 self-employed are owners of unincorporated businesses.

Table 4: Random-Effects Probit Model of Access to Credit

Credit Products				
	Credit Card, t	**Car Loan, $t = 2003$**	**Charge Card, t**	**Refinance Mortgage, t**
	(1)	(2)	(3)	(4)
CMA, $t-1$	1.081***	1.047***	1.047*	1.050**
	(0.032)	(0.024)	(0.029)	(0.018)
Observations	17,279	3,003	19,097	5,547
Reduction in Access to Credit				
	Application for new credit has been denied, t	**Reduction in the Available Limit for Credit Cards, t**	**Being Asked to Pay Off Remaining Balance, t**	**Filing Bankruptcy, t**
	(5)	(6)	(7)	(8)
CMA, $t-1$	0.991*	0.952*	0.862**	0.913*
	(0.036)	(0.025)	(0.059)	(0.048)
Observations	14,283	6,080	5,200	13,283

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes. Table presents the relative risk ratios from the random-effect probit model of access to credit market equation (7), where the random-effect term $\nu_i \in N(0, \sigma_\nu^2)$. The dependent variables are dummy variables that equal 1 if the respondent applied for a new credit card, applied for a car loan, opened a new charge card, refinanced a mortgaga, had an application for new credit denied, had available limit on a credit card reduced, was asked to pay off the remaining loan balance, or filed for bankruptcy in the last 12 months. Columns present the full estimates of eq.(7) with correlated random effects, using the Mundlak-Chamberlain devise to avoid a problem of possible correlation of unobserved individual permanent heterogeneity with explanatory variables. Other explanatory variables include the value of large durable assets, durable logarithm of differences in average wages of salary workers and unincorporated self-employed, individual characteristics such as age, gender, etc., local economic conditions, year dummies, four regional dummies, fixed effects for the first year of the stochastic process, and the Mundlack device.

Table 5: Dynamic Multinomial Logit Model of Employment Choices.

The types of self-employment are based on the legal status of the business (main specification)

	$\eta_{ij} \in N(0;1)$			The discrete factor approximation of η_{ij}		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$
T-1 SE, t	0.191***	50.034***	12.107***	0.185***	74.910***	17.372***
	(0.074)	(28.507)	(7.376)	(0.066)	(27.314)	(6.502)
T-2 SE, t	0.181***	13.791***	23.329***	0.185***	21.695***	48.681***
	(0.059)	(6.212)	(11.985)	(0.051)	(6.813)	(13.917)
NE, t	0.040***	0.390**	0.223***	0.031***	0.276***	0.172***
	(0.007)	(0.148)	(0.087)	(0.003)	(0.080)	(0.047)
CMA	1.182**	1.323*	0.974	1.116**	1.213**	0.981
	(0.090)	(0.221)	(0.172)	(0.053)	(0.111)	(0.088)
Assets	1.095*	1.070**	0.987	1.235	1.528***	1.187**
	(0.062)	(0.119)	(0.121)	(0.849)	(0.110)	(0.083)
log (w_{PE})- log (w_{SE_2})	1.435***	1.751	0.756**	1.316***	1.662	0.928**
	(0.335)	(0.690)	(0.305)	(0.044)	(0.592)	(0.014)
Control variables, X_i and X_{it}						
Age	1.250***	1.311***	1.310***	1.281***	1.342***	1.346***
	(0.039)	(0.116)	(0.085)	(0.045)	(0.109)	(0.097)
Age sq.	0.997***	0.996***	0.997***	0.997***	0.996***	0.996***
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Female	0.930	0.547***	0.719*	0.912	0.536***	0.705**
	(0.089)	(0.100)	(0.125)	(0.092)	(0.099)	(0.122)
Race:						
Black	0.749***	0.698	0.575***	0.751***	0.698*	0.572***
	(0.074)	(0.154)	(0.121)	(0.077)	(0.152)	(0.116)
Other	0.670**	0.380***	0.832	0.651**	0.369***	0.810
	(0.107)	(0.122)	(0.256)	(0.112)	(0.127)	(0.238)
Hispanic	0.851	1.290	0.569*	0.852	1.283	0.565*
	(0.146)	(0.444)	(0.171)	(0.147)	(0.408)	(0.181)
Married	1.013	0.952	1.283	1.015	0.951	1.286
	(0.131)	(0.282)	(0.328)	(0.135)	(0.269)	(0.334)

	$\eta_{ij} \in N(0;1)$			The discrete factor approximation of η_{ij}		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$
Control variables, X_i and X_{it}						
Education:						
Some college	1.163*	1.427*	1.423*	1.174*	1.441*	1.434*
	(0.104)	(0.303)	(0.257)	(0.112)	(0.307)	(0.276)
Bachelor's degree	0.795	1.550	0.662	0.814	1.599	0.688
	(0.189)	(0.769)	(0.384)	(0.197)	(0.743)	(0.341)
Grad degree	1.363**	2.344***	1.806**	1.427**	2.455***	1.878**
	(0.186)	(0.640)	(0.456)	(0.217)	(0.671)	(0.483)
Number of kids:						
1 child	1.060	0.812	0.937	1.061	0.815	0.938
	(0.120)	(0.203)	(0.204)	(0.134)	(0.207)	(0.215)
Fewer than 4 kids	0.756**	0.835	0.676*	0.748**	0.827	0.668*
	(0.091)	(0.197)	(0.145)	(0.094)	(0.190)	(0.149)
More than 4 kids	0.566	0.874	0.849	0.583	0.900	0.870
	(0.252)	(0.571)	(0.438)	(0.246)	(0.563)	(0.514)
Number of HH members:						
2 adults in HH	1.283**	1.539	0.861	1.314**	1.581	0.883
	(0.146)	(0.490)	(0.210)	(0.165)	(0.461)	(0.229)
Fewer than 4 adults in HH	1.359**	1.662	1.015	1.398**	1.721	1.049
	(0.205)	(0.641)	(0.292)	(0.218)	(0.599)	(0.323)
More than 4 adults in HH	1.515	1.959	1.850	1.510	1.951	1.840
	(0.555)	(1.604)	(1.459)	(0.618)	(1.757)	(1.310)
Local Economic Conditions, L_{it}	Yes	Yes	Yes	Yes	Yes	Yes
Time-averaged variables $\overline{X_i^+}$, $\overline{Z_i^+}$, $\overline{L_i^+}$	Yes	Yes	Yes	Yes	Yes	Yes
Initial conditions, Y_{i1}, X_{i1} and L_{i1}	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Var(η_{ij})	0.121	0.110	0.413	0.072	0.042	0.034
	(0.019)	(0.022)	(0.021)	(0.029)	(0.028)	(0.026)
Cov $(\eta_{i,PE};\eta_{i,SE_1})$		0.211			0.031	
		(0.013)			(0.023)	
Cov $(\eta_{i,SE_1};\eta_{i,SE_2})$		0.277			0.037	
		(0.013)			(0.023)	
Cov $(\eta_{i,PE};\eta_{i,SE_2})$		0.141			0.037	
		(0.016)			(0.022)	
Observations		11,922			11,922	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

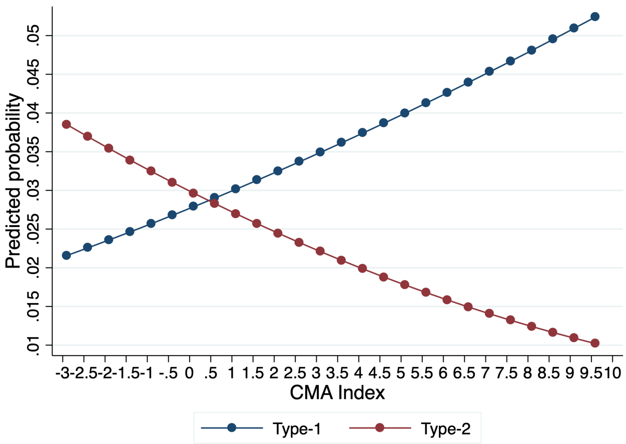
Notes. Table presents the relative risk ratios from the dynamic multinomial logit model of employment with correlated random effects. The dependent variable is employment status in t+1, with the non-employed chosen as the base outcome. Columns 1 and 2 present the full estimates of eq.4 with correlated random effects, using the WRS solution to the endogeneity of initial conditions. The following variables are included but not shown: year dummies, four regions, fixed effects for the first year of the stochastic process, the Mundlack device and the intercept, regional economic characteristics. The omitted category of race is white, of number of kinds is no kids, of number of adults is 1 adult .

Table 6: Post-estimation predictions: Average marginal effect of CMA on transition probability

	Type-1 SE, $t+1$	Type-2 SE, $t+1$	PE, $t+1$	NE, $t+1$
Sector size in $t+1$	0.22 pp	-0.20 pp	0.71 pp	-0.73 pp
	(0.0005)	(0.0008)	(0.0021)	(0.0035)
Transition probabilities				
$PE_t \rightarrow$	0.08 pp	-0.07 pp	0.67 pp	-0.66 pp
	(0.0002)	(0.0003)	(0.0024)	(0.0025)
Type-1 $\text{SE}_t \rightarrow$	3.7 pp	-1.5 pp	-1.6 pp	-0.6 pp
	(0.0031)	(0.0043)	(0.0030)	(0.0029)
Type-2 $\text{SE}_t \rightarrow$	1.8 pp	-5.1 pp	3.8 pp	-0.5 pp
	(0.764)	(0.0078)	(0.0056)	(0.0021)
$NE_t \rightarrow$	0.3 pp	-1.2 pp	3.3 pp	-2.4 pp
	(0.0012)	(0.0032)	(0.0082)	(0.0091)

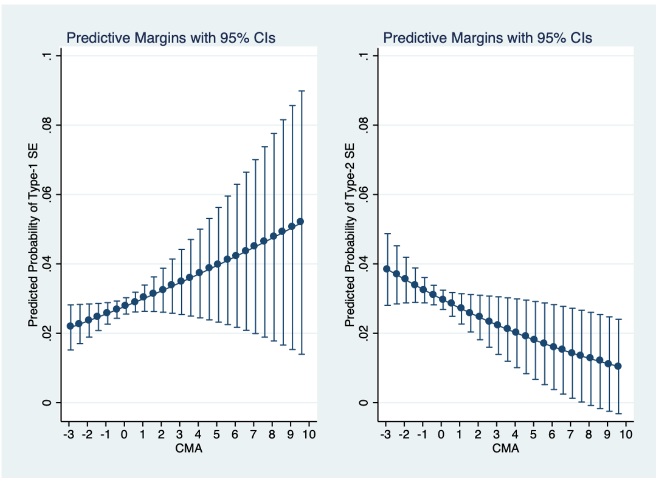
Notes. This table shows the average marginal effect of the increase in credit market accessibility by one standard deviation on the transition probability to self-employment. Types of self-employed are defined based on the type of business. Type-1 self-employed are owners of incorporated business, Type-1 self-employed are owners of unincorporated business

Figure 8: Marginal effects of CMA on the predicted probability of being SE



Notes. Figure shows the predicted probabilities of being in one of the two type of self-employed at different values of the credit market accessibility index. The index is standardized to have zero mean and a standard deviation of one. Predictions are obtained at sample means of covariates and based on the WRS model reported in Table 5.

Figure 9: Marginal effects of CMA on the predicted probability of being SE



Notes. Figure shows the predicted probabilities of being in one of the two type of self-employed at different values of the credit market accessibility index. The index is standardized to have zero mean and a standard deviation of one. Predictions are obtained at sample means of covariates and based on the WRS model reported in Table 5.

Table 7: Results of Dynamic Multinomial Logit Model of Employment Choices for Different Definitions of Types of Self-employed

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Based on type of business			Entrepreneurial abilities			Aggregated characteristics			EITC		
VARIABLES	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$
T-1 SE, t	0.189***	76.545***	17.770***	0.132***	47.770***	6.969***	0.325**	129.637***	18.958***	0.186***	71.827***	17.290***
	(0.062)	(26.515)	(6.088)	(0.038)	(15.098)	(2.778)	(0.172)	(68.292)	(10.526)	(0.067)	(28.283)	(7.026)
T-2 SE, t	0.191***	22.426***	50.420***	0.232***	22.060***	75.438***	0.139***	9.037***	47.405***	0.175***	17.667***	39.541***
	(0.054)	(7.104)	(16.034)	(0.081)	(8.014)	(28.076)	(0.034)	(2.940)	(13.542)	(0.050)	(6.099)	(12.151)
NE, t	0.027***	0.241***	0.150***	0.028***	0.194***	0.169***	0.028***	0.190***	0.182***	0.027***	0.171***	0.208***
	(0.003)	(0.072)	(0.040)	(0.003)	(0.053)	(0.053)	(0.003)	(0.062)	(0.048)	(0.003)	(0.060)	(0.051)
CMA	1.103**	1.197**	0.969	1.103**	1.070**	1.100	1.105**	1.042**	1.108	1.108**	1.204**	0.981*
	(0.051)	(0.105)	(0.087)	(0.054)	(0.048)	(0.112)	(0.054)	(0.043)	(0.093)	(0.051)	(0.113)	(0.045)
Observations	12,546			11,930			11,930			12,433		

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes. Table presents the relative risk ratios from the dynamic multinomial logit model of employment with correlated random effects eq.(5). The dependent variable is employment status in $(t+1)$, the non-employment status is chosen as the base outcome. Columns present the full estimates of eq.(5) with correlated random effects, using the WRS solution to the endogeneity of initial conditions. The following variables are included but not shown: the value of large durable assets, durable logarithm of differences in average wages of salary workers and unincorporated self-employed, individual characteristics, local economic conditions, year dummies, four regions, fixed effects for the first year of the stochastic process, and the Mundlak-Chamberlain device. The omitted category of employment status at period t is paid employment.

Table 8: Results of Dynamic Multinomial Logit Model of Employment Choices for Different Identifying Assumptions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Exogenous CMA			Lag of CMA			Individuals who did not move			Local economic conditions		
VARIABLES	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$	PE, $t+1$	T-1 SE, $t+1$	T-2 SE, $t+1$
T-1 SE, t	0.189***	76.545***	17.770***	0.180***	75.619***	17.809***	0.108***	97.850***	17.100***	0.191***	50.034***	12.107***
	(0.062)	(26.515)	(6.088)	(0.059)	(26.515)	(6.085)	(0.044)	(46.723)	(7.290)	(0.074)	(28.507)	(7.376)
T-2 SE, t	0.191***	22.426***	50.420***	0.183***	23.598***	52.766***	0.205***	34.496***	51.676***	0.181***	13.791***	23.329***
	(0.054)	(7.104)	(16.034)	(0.053)	(7.642)	(17.014)	(0.076)	(15.043)	(20.958)	(0.059)	(6.212)	(11.985)
NE, t	0.027***	0.241***	0.150***	0.027***	0.236***	0.143***	0.025***	0.333**	0.121***	0.031***	0.276***	0.172***
	(0.003)	(0.072)	(0.040)	(0.003)	(0.074)	(0.040)	(0.004)	(0.151)	(0.046)	(0.003)	(0.080)	(0.047)
CMA	1.103**	1.197**	0.969	1.100**	1.205**	0.986	1.075	1.242*	0.950	1.182**	1.323*	0.974
	(0.051)	(0.105)	(0.087)	(0.053)	(0.111)	(0.089)	(0.074)	(0.144)	(0.126)	(0.062)	(0.119)	(0.121)
Observations	12,546			10,877			6,359			11,922		

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes. Table presents the relative risk ratios from the dynamic multinomial logit model of employment with correlated random effects. The dependent variable is employment status in $(t+1)$, the non-employment status is chosen as the base outcome. Columns (1)-(3) present the full estimates of the base model (eq.(4)) with correlated random effects under the assumption of exogeneity of CMA, using the WRS solution to the endogeneity of initial conditions. Columns (4)-(6) present the full estimates of eq.(5) with correlated random effects using a lag of CMA measure, including local economic conditions and using the WRS solution to the endogeneity of initial conditions. Columns (7)-(9) present the full estimates of eq.(5) with correlated random effects only for individuals who did not move from their original place of living, including local economic conditions and using the WRS solution to the endogeneity of initial conditions. Columns (10)-(12) present the full estimates of eq.(5) with correlated random effects, including local economic conditions, withing-means of CMA and using the WRS solution to the endogeneity of initial conditions. The following variables are included but not shown: the value of large durable assets, durable logarithm of differences in average wages of salary workers and unincorporated self-employed, individual characteristics, local economic conditions, year dummies, four regions, fixed effects for the first year of the stochastic process, and the Mundlak-Chamberlain device. The omitted category of employment status at period t is paid employment.

Table 9: Results of Dynamic Multinomial Logit Model of Employment Choices with Endogenous CMA

VARIABLES	(1) PE, $t+1$	(2) T-1 SE, $t+1$	(3) T-2 SE, $t+1$	(4) CMA
T-1 SE, t	0.179***	55.906***	16.721***	
	(0.076)	(24.659)	(7.263)	
T-2 SE, t	0.162***	17.028***	34.201***	
	(0.057)	(6.094)	(10.662)	
NE, t	0.031***	0.238***	0.146***	
	(0.004)	(0.079)	(0.045)	
CMA	1.119*	1.333**	0.963	
	(0.042)	(0.115)	(0.163)	
IV				0.801***
				(0.034)
Observations	11,922			

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes. Table presents the relative risk ratios from the joint estimation of the individual i labor decisions (the dynamic multinomial logit model of employment with correlated random effects) and CMA (eq. 6). The dependent variable is employment status in $(t+1)$ and CMA in (t). The non-employment status is chosen as the base outcome in the dynamic multinomial logit model. Variable "IV" equals to 1 if the individual i lives in the Census tract where two large banks undergo a merger in year t. The following variables are included but not shown: the value of large durable assets, durable logarithm of differences in average wages of salary workers and unincorporated self-employed, individual characteristics, local economic conditions, year dummies, four regions, fixed effects for the first year of the stochastic process, and the Mundlak-Chamberlain device. The omitted category of employment status at period t is paid employment.

Table 10: Policy Simulations

VARIABLES	PE, before	PE, after	T-1 SE, before	T-1 SE, after	Change in T-1 SE share	T-2 SE, before	T-2 SE, after	Change in T-2 SE share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Marginal effects for components of CMA index								
Aver. distance to 10 nearest banks, miles	-0.00036		-0.00003			0.0005**		
	(0.001)		(0.0002)			(0.0003)		
Number of banks within 5 miles	0.00034**		0.00008			-0.0001		
	(0.0002)		(0.00008)			(0.00011)		
Number of bank workers per 1000 pop.	0.005		0.0009			-0.004		
	(0.008)		(0.003)			(0.004)		
Distance to the nearest bank								
Policy 1: open 37 bank offices withing 5 miles	0.820	0.834	0.027	0.031	+0.004	0.035	0.032	-0.003
if there were no banks within 5 miles	(0.007)	(0.005)	(0.003)	(0.002)		(0.0003)	(0.0002)	
Policy 2: use online banking	0.835	0.927	0.0298	0.0387	+0.0089	0.0333	0.0023	-0.031
(average distance $\rightarrow$ 0)	(0.007)	(0.004)	(0.001)	(0.003)		(0.001)	(0.003)	
Number of bank workers								
Policy 3: increase the number of bank workers	0.840	0.841	0.0308	0.0312	+0.0004	0.0302	0.0291	-0.0011
from 2 per 10,000 pop. to 5 per 10,000 pop.	(0.005)	(0.004)	(0.002)	(0.003)		(0.002)	(0.001)	

Notes. Table show a marginal effect for components of CMA index, also table shows the predicted probabilities of being paid-employed, Type-1 self-employed and Type-2 self-employed before and after a given policy is implemented. The predicted probability of being non-employed is $1 - \widehat{Pr_{PE}} - \widehat{Pr_{SE_1}} - \widehat{Pr_{SE_2}}$ (not shown). The marginal effects of policies on the employment composition are predicted based on the main model (eq.5) estimates shown in Table 5. Predicted probabilities are evaluated at the starting values indicated in the table and at sample means of all other covariates. The standard errors are estimated using the delta method and reported in parentheses. Columns (5) and (8) shows the change in the relative size of sectors due to a simulated policy.

10 Appendix

Solution of the Theoretical Model

In this model, time is discrete. In contrast to Evans and Jovanovic (1989) and Levine and Rubinstein (2018), I abstract to a stylized two-period model for the clarity of predictions, although the extension to the infinite horizon is straightforward. Agents live for two periods - the present (t) and the future ($t+1$)[21]. Each period the agent can be in one of three labor states: Type-1 self-employed ($j = SE_1$), Type-2 self-employed ($j = SE_2$) or paid employed worker ($j = PE$). The set of all possible choices for the agent:

$$\begin{aligned} J =& \{(PE_t, PE_{t+1}); (PE_t, SE_{1t+1}); (PE_t, SE_{2t+1}); (SE_{1t}, PE_{t+1}); (SE_{1t}, SE_{1t+1}); \\ & (SE_{1t}, SE_{2t+1}); (SE_{2t}, PE_{t+1}); (SE_{2t}, SE_{1t+1}); (SE_{2t}, SE_{2t+1})\} \end{aligned}$$

Each period employment type-specific uncertainty (ϵ_t^j and ϵ_{t+1}^j) realizes. For simplicity in describing the model, assume that the error terms are known at the beginning of the first period so that the optimization problem may be started in terms of individual making all decisions at the beginning of the first period, so at the begin of the period t, the agent makes the static labor decision $j \in J$ for both periods. If instead the random variable for period $t+1$ was not known until the beginning of that period, then the model could be extended with the decisions in period t based on the distribution of the error term in period $t+1$. Assuming that the error term for the second period is not realized until after the first period decisions are made complicates the notation but do not change the nature of the problem or solution.

Individuals select in which labor market sector to participate based on their productivity (ability), θ, and the amount of available assets, a_t. The agent is allowed to take a business loan only if her amount of assets exceeds the value a_b that is determined by the bank (see Fig.10). Individuals with access to financial markets ($a_t > a_b$) may borrow to finance some of their capital acquisitions; others must finance their purchases from their beginning-of-period assets. At the end of the period, individuals (a) receive revenue

[21] Period t is the beginning of the agent's life, and she did not participate in labor market at period $t-1$.

from any business operations, paid-employment, and financial investments, (b) repay any loans and accrued interest, (c) pay switching cost in period $t+1$ if they changed labor market sectors from period t, (d) receive interest on savings, and (e) purchase the consumption goods. The agent is permitted to borrow at the period t against the next period's revenues at the period $t+1$, and any assets remaining at the end of the period t, a_{t+1}, are carried forward to the next period, $t+1$.

The individual selects the labor market sectors, her borrowing to purchase capital, and her consumption to maximize her expected discounted utility subject to the given below constraints. The discounted utility for an individual who selects employment types j and who consumes c_t and c_{t+1} :

$$\max_{c_t \geq 0,\ c_{t+1} \geq 0} V = U^j(c_t) + \beta U^j(c_{t+1}) + \epsilon_t^j + +\epsilon_{t+1}^j$$

where ϵ^j is the choice-specific error term.

The decision rule $\delta(j)$:

$$\delta(j) = \arg\max_j \left(U^j(c_t) + \beta U^j(c_{t+1}) + \epsilon_t^j + \epsilon_{t+1}^j \right)$$

Given this set-up, the budget constraints for a two-period version of the model can be most easily derived by first determining the constraint involving the assets at the end of period t, as dependent upon the decisions in that period, and then determining the constraint involving consumption in period $t+1$, as dependent upon the assets carried forward to the period and the decisions made during the period. For any labor market choice, the assets at the end of the period t consist of any revenue from business operations, paid-employment, and interest on financial investments minus (a) consumption expenditures and (b) any repayment of business borrowings and the accrued interest, and the transaction costs of borrowing.

The budget constraint of self-employed worker (any type) at period t:

$$a_{t+1}+c_t = f(\theta, k_t) - r(k_t - a_t)\mathbb{1}((k_t - a_t) \geq 0) - (k_t - a_t)\mathbb{1}((k_t - a_t) < 0) - \psi_t \mathbb{1}((k_t - a_t) \geq 0)$$

where c_t is consumption in period t, $f(\theta, k_t)$- the production function, k_t - the amount of capital invested in the business at period t, a_{t+1} - savings in period t, r - the gross cost of capital (one plus the interest rate), θ - agent's productivity level. If $a_t \leq k_t$ the agent is a net-borrower, and $r(k_t - a_t)$ is the amount she repays at the end of the period; if $(a_t \geq k_t)$ she is a net-saver. The model also includes the non-interest cost of borrowing, ψ_t, that may contain not only the direct fees charged at closing such as origination and application fees but also the indirect cost to borrowers such as the value of time spent preparing application documents and traveling to the bank.

The paid-employed worker is allowed to take a one-period non-business loan (e.g., consumer loan, mortgage) at a gross interest rate r_b (one plus the interest rate). The budget constraint of the paid-employed worker at period t:

$$a_{t+1} + c_t = \theta w_t + a_t + b_t(1 - r) - \psi_t \mathbb{1}(b_t \neq 0)$$

where c_t is consumption in period t, a_t - initial wealth at period t, a_{t+1} - savings in period t, w_t is the wage at period t, θ - agent's productivity level, b_t - a one-period non-business loan, r - the gross cost of loan (one plus the interest rate), ψ_t - non-interest cost of borrowing.

In the second period, the individual has no incentive to carry assets forward to the next period, so consumption equals the revenue from all sources minus (a) any repayment of business borrowing, accrued interest, and the transaction cost of borrowing, and (b) any switching cost as a result of a change from the type of labor market from the first period. If at period t the agent was a paid-worker, she pays the switching cost π_t^{PE} to any type of self-employed at period $t + 1$. If she worked as Type-2 self-employed at period t, she pays the switching cost $\pi_t^{SE_1}$ to Type-1 self-employed at period $t + 1$.

The budget constraint of self-employed worker (any type) at period $t + 1$:

$$c_{t+1} = f(\theta, k_{t+1}) - r(k_{t+1} - ra_{t+1})\mathbb{1}\left((k_{t+1} - ra_{t+1}) \geq 0\right) - (k_{t+1} - ra_{t+1})\mathbb{1}\left((k_{t+1} - ra_{t+1}) < 0\right) -$$
$$- \psi_{t+1}\mathbb{1}\left((k_{t+1} - ra_{t+1}) \geq 0\right) - \pi_{t+1}^{PE}\mathbb{1}\left(y_t = y^{PE}\right) - \pi_{t+1}^{SE_1}\mathbb{1}\left((y_t = y^{SE_2}) \;\&\; (y_{t+1} = y^{SE_1})\right)$$

where c_{t+1} - consumption in period $t + 1$, $f(\theta, k_{t+1})$- the production function, k_{t+1} - the amount of capital invested in the business at period $t+1$, a_t - initial wealth at period t, a_{t+1} - savings in period t, r - the gross cost of capital (one plus the interest rate), ra_{t+1} - wealth in period $t+1$, θ - agent's productivity level. If $ra_{t+1} \leq k_{t+1}$, the agent is a net-borrower, and $r(k_{t+1} - ra_{t+1}))$ is the amount she repays at the end of the period; if $a_{t+1} \geq k_{t+1}$, she is a net-saver. If the agent decides to switch from paid-employment to other labor states, she has to pay the cost of switching from paid employment, π^{PE}_{t+1} that may include the cost of the license, security deposit for office, cost of time invested in new skills. Being a Type-1 self-employed requires higher entrepreneurial skills than Type-2 self-employed, it means if the agent switches to Type-1 self-employment, she has to pay the cost of switching, $\pi^{SE_1}_{t+1}$, that includes the cost of time invested in new skills.

The budget constraint of the paid-employed worker at period $t+1$:

$$c_{t+1} = \theta w_{t+1} + ra_{t+1} + b_{t+1}(1-r) - \psi_{t+1}\mathbb{1}(b_{t+1} \neq 0)$$

where c_{t+1} - consumption in period $t+1$, a_{t+1} - savings in period t, r - the gross cost of capital (one plus the interest rate), ra_{t+1} - wealth in period $t+1$, w_{t+1} - the wage at period $t+1$, θ - agent's productivity level, b_{t+1} - a one-period non-business loan, ψ_{t+1}-non-interest cost of borrowing.

Given the additive nature of the business revenue and any loan costs in the constraints, the optimal capital decisions can be made separately from the consumption decisions. In either period, a self-employed individual with access to the capital market selects the amount of capital that maximizes her business profit or:

$$\begin{aligned}
\text{Period } t\text{:} \quad & \max_{0 \leq k_t \leq \lambda a_t} [f(\theta, k_t) - r(k_t - a_t)\mathbb{1}\left((k_t - a_t) \geq 0\right) - (k_t - a_t)\mathbb{1}\left((k_t - a_t) < 0\right)] \\
\text{Period } t+1\text{:} \quad & \max_{0 \leq k_{t+1} \leq \lambda r a_{t+1}} [f(\theta, k_{t+1}) - r(k_{t+1} - ra_{t+1})\mathbb{1}\left((k_{t+1} - ra_{t+1}) \geq 0\right) - \\
& - (k_{t+1} - ra_{t+1})\mathbb{1}(k_{t+1} - ra_{t+1}) < 0)]
\end{aligned}$$

Assuming that the production function is strictly concave in the capital and making the Inada assumption that the marginal product of capital is infinity at zero, the optimal level of capital is positive and finite, and is denoted k^*. If a self-employed individual (Type-2) does not have access to the capital markets and cannot finance the optimal capital with her own assets, then she uses all of her assets to purchase capital. The resulting revenue is greater than if she had used only part of the assets to purchase capital and had put the remainder in a financial instrument earning a gross return r.

As in Evans and Jovanovic (1989), I assume that the agent can borrow an amount that is proportional to her wealth ($k_t \leq \lambda a_t$ and $k_{t+1} \leq \lambda r a_{t+1}$, where $\lambda \geq 1$). If the amount of wealth exceeds the amount required to finance ($\lambda a_t \geq k_t^*$ and $\lambda r a_{t+1} \geq k_{t+1}^*$), the remaining wealth is invested at the rate r, and it's "unconstrained case". If the optimal amount of capital is lower than the amount of wealth ($\lambda a_t \leq k_t^*$ and $\lambda r a_{t+1} \leq k_{t+1}^*$), the agent borrows the maximum available amount of loan, and it is "constrained case".

Agents sort into employment type j to maximize her utility:

$$V = U^j(c_t) + \beta U^j(c_{t+1}) + \epsilon_t^j + \epsilon_{t+1}^j$$

where ϵ^j is the choice-specific error term. All possible labor choices: $J = \{(PE_t, PE_{t+1});$ $(PE_t, SE_{1t+1}); (PE_t, SE_{2t+1}); (SE_{1t}, PE_{t+1}); (SE_{1t}, SE_{1t+1}); (SE_{1t}, SE_{2t+1}); (SE_{2t}, PE_{t+1});$ $(SE_{2t}, SE_{1t+1}); (SE_{2t}, SE_{2t+1})\}$

The decision rule $\delta(j)$:

$$\delta(j) = \arg\max_j \left(U^j(c_t) + \beta U^j(c_{t+1}) + \epsilon_t^j + \epsilon_{t+1}^j\right)$$

Assume that $U(c_t) = log(c_t)$ and $\epsilon^j \in N(0; 1)$ for simplicity.

If the agent works as a paid-employed worker, her budget constraints:

$$\begin{aligned} &\text{Period } t\text{:} \quad a_{t+1} + c_t = \theta w_t + a_t + b_t(1 - r_b) - \psi_t \\ &\text{Period } t+1\text{:} \quad c_{t+1} = \theta w_{t+1} + r a_{t+1} + b_{t+1}(1 - r_b) - \psi_{t+1} \end{aligned}$$

where c_t is consumption in period t, c_{t+1} - consumption in period $t+1$, a_t - initial wealth at period t, a_{t+1} - savings in period t, r - the gross cost of capital (one plus the interest rate), $r a_{t+1}$ - wealth in period $t+1$, w_t is the wage at period t, w_{t+1} - the wage at period $t+1$, θ - agent's productivity level, b_t - a one-period non-business loan at period t, b_{t+1} - a one-period non-business loan at period $t+1$, r_b - an gross interest rate of non-business loan, ψ_t - non-interest cost of borrowing at period t, ψ_{t+1} - non-interest cost of borrowing at period $t+1$.

If the agent works as a Type-1 self-employed worker (only if $a_t \geq a_b, a_{t+1} \geq a_b$), she can borrow money from the bank. In this case she maximizes her production function ($f(\theta, k_t) = \theta k_t^\alpha$, where $\alpha < 1$) making decision about amount of capital invested in

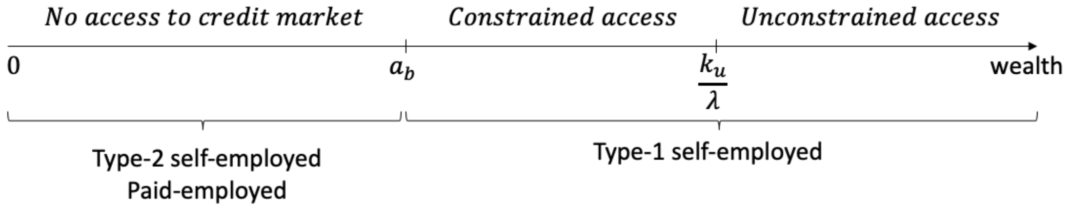


Figure 10: Wealth and The Access to Credit Market

the business, and she can use her wealth as a collateral. In this case agent's budget constraints:

$$\text{Period } t: \quad a_{t+1} + c_t = \max_{0 \le k_t \le \lambda a_t} [\theta (k_t)^\alpha - r(k_t - a_t)\mathbb{1}((k_t - a_t) \ge 0) - (k_t - a_t)\mathbb{1}((k_t - a_t) < 0) -$$

$$- \psi_t \mathbb{1}((k_t - a_t) \ge 0)]$$

$$\text{Period } t+1: \quad c_{t+1} = \max_{0 \le k_{t+1} \le \lambda r a_{t+1}} [\theta (k_{t+1})^\alpha - r(k_{t+1} - r a_{t+1})\mathbb{1}((k_{t+1} - r a_{t+1}) \ge 0) - (k_{t+1} - r a_{t+1}) \cdot$$

$$\cdot \mathbb{1}((k_{t+1} - r a_{t+1}) < 0) - \psi_{t+1}\mathbb{1}((k_{t+1} - r a_{t+1}) \ge 0) - \pi_{t+1}^{PE}\mathbb{1}(y_t = y^{PE}) - \pi_{t+1}^{SE_1}\mathbb{1}(y_t = y^{SE_2})]$$

where c_t is consumption in period t, c_{t+1} - consumption in period $t+1$, k_t - the amount of capital invested in the business at period t, k_{t+1} - the amount of capital invested in the business at period $t+1$, a_t - initial wealth at period t, a_{t+1} - savings in period t, r - the gross cost of capital (one plus the interest rate), ra_{t+1} - wealth in period $t+1$, θ - agent's productivity level. If $a_t \le k_t$ or $ra_{t+1} \le k_{t+1}$, the agent is a net-borrower, and $r(k_t - a_t)$ or $r(k_{t+1} - ra_{t+1}))$ is the amount she repays at the end of the period. If $(a_t \ge k_t)$ or $(a_{t+1} \ge k_{t+1})$, she is a net-saver, and $(k_t - ra_t)$ is the amount she has as wealth at the period $t+1$. If at period t the agent was a paid-worker, she pays switching cost π_t^{PE} at period $t+1$. If she worked as Type-2 self-employed at period t, she pays switching cost $\pi_t^{SE_1}$ at period $t+1$. As in Evans and Jovanovic (1989), I assume that the agent can borrow an amount that is proportional to her wealth ($k_t \le \lambda a_t$ and $k_{t+1} \le \lambda r a_{t+1}$, where $\lambda \ge 1$). If the amount of wealth exceeds the amount required to finance ($\lambda a_t \ge k_t^*$ and $\lambda r a_{t+1} \ge k_{t+1}^*$), the remaining wealth is invested at the rate r, and it's "unconstrained case". If the optimal amount of capital is lower than the amount of wealth ($\lambda a_t \leq k_t^*$ and $\lambda r a_{t+1} \leq k_{t+1}^*$), the agent borrows the maximum available amount of loan, and it is "constrained case".

In constrained case $k_t^* = \lambda a_t$ and $k_{t+1}^* = \lambda r a_{t+1}$. In unconstrained case the amount of capital invested in the business is higher than the amount of capital in constrained case: $k_t^* = (\frac{r}{\theta})^{\frac{\alpha}{\alpha-1}}$ and $k_{t+1}^* = (\frac{r}{\theta})^{\frac{\alpha}{\alpha-1}}$. And if the agent uses her personal savings $k_t \leq a_t$ and it is lower than the amount of capital in contained case, therefore the agent always prefers to borrow, instead of paying out of pocket.

Budget constraints of unconstrained Type-1 self-employed worker, where $\gamma = \left(\frac{r}{\theta}\right)^{\frac{1}{\alpha-1}}$:

$$\text{Period } t\text{:} \quad a_{t+1} + c_t = \theta\gamma^\alpha - r\left(\gamma - a_t\right) - \psi_t$$

$$\text{Period } t+1\text{:} \quad c_{t+1} = \theta\gamma^\alpha - r\left(\gamma - r a_{t+1}\right) - \psi_{t+1} - \pi_{t+1}^{PE}\mathbb{1}(y_t = y^{PE}) - \pi_{t+1}^{SE_1}\mathbb{1}(y_t = y^{SE_2})$$

Budget constraints of constrained Type-1 self-employed worker:

$$\text{Period } t\text{:} \quad a_{t+1} + c_t = \theta(\lambda a_t)^\alpha - r a_t(\lambda - 1) - \psi_t$$

$$\text{Period } t+1\text{:} \quad c_{t+1} = \theta(\lambda r a_{t+1})^\alpha - r a_{t+1}(\lambda - 1) - \psi_{t+1} - \pi_{t+1}^{PE}\mathbb{1}(y_t = y^{PE}) - \pi_{t+1}^{SE_1}\mathbb{1}(y_t = y^{SE_2})$$

Further analysis will investigate the role of the non-interest costs of borrowing in labor market choices, and for the clarity of predictions I will consider only unconstrained case, but although the extension to the constrained case is straightforward.

If the agent works as a Type-2 self-employed worker (if $a_t \leq a_b, a_{t+1} \leq a_b$), she does not have access to credit market. In this case she maximizes her production function making decision about amount of capital invested in the business, but she uses assets as a capital. In this case agent's budget constraints:

$$\text{Period } t\text{:} \quad a_{t+1} + c_t = \max_{k_t \leq a_t}[\theta(k_t)^\alpha]$$

$$\text{Period } t+1\text{:} \quad c_{t+1} = \max_{k_{t+1} \leq r a_{t+1}}[\theta(k_t)^\alpha] - \pi_{t+1}^{PE}\mathbb{1}(y_t = y^{PE})]$$

Budget constraints of Type-2 self-employed worker:

$$\text{Period } t\text{:} \quad a_{t+1} + c_t = \theta(a_t)^\alpha$$

$$\text{Period } t+1\text{:} \quad c_{t+1} = \theta(r a_{t+1})^\alpha - \pi_{t+1}^{PE}\mathbb{1}(y_t = y^{PE})$$

Utility Functions for All Possible Labor Choices

- $\{SE_{1t}, SE_{1t+1}\}$

$$U^{(SE_{1t},SE_{1t+1})} - \epsilon^{(SE_{1t},SE_{1t+1})} = (1+\beta)log\left(\frac{(r^2+1)(\theta\gamma^\alpha - r\gamma) + r^3 a_t - \psi_{t+1} - r^2\psi_t}{r^2+r}\right) + \beta log(r)$$

- $\{SE_{1t}, PE_{t+1}\}$

$$U^{(SE_{1t},PE_{t+1})} - \epsilon^{(SE_{1t},PE_{t+1})} = (1+\beta)log\left(\frac{r\theta\gamma^\alpha - r^2(\gamma - a_t) - r\psi_t + \theta w_{t+1} + b_{t+1}(1-r_b) + \psi_{t+1}}{(r+1)}\right)$$

- $\{SE_{1t}, SE_{2t+1}\}$

$$U^{(SE_{1t},SE_{2t+1})} - \epsilon^{(SE_{1t},SE_{2t+1})} = (\alpha+\beta)log\frac{(\theta\gamma^\alpha - r(\gamma - a_t) - \psi_t)}{r+\alpha} + (\alpha+\beta)log(r) + log(\alpha^\alpha\theta)$$

- $\{PE_t, PE_{t+1}\}$

$$U^{(PE_t,PE_{t+1})} - \epsilon^{(PE_t,PE_{t+1})} = (1+\beta)log\left(\frac{r\theta w_t + r a_t + r b_t(1-r_b) + \theta w_{t+1} + b_{t+1}(1-r_b)}{r+1}\right)$$

- $\{PE_t, SE_{1t+1}\}$

$$U^{(PE_t,SE_{1t+1})} - \epsilon^{(PE_t,SE_{1t+1})} = (1+\beta)log\left(\frac{r^2\theta w_t + r^2 a_t + r^2(b_t(1-r_b) - \psi_t) + \theta\gamma^\alpha - r\gamma - \psi_{t+1} - \pi_{t+1}^{PE}}{(r^2+r)}\right) + \beta log(r)$$

- $\{PE_t, SE_{2t+1}\}$ (under assumption $\pi_{t+1}^{PE} = 0$[22])

$$U^{(PE_t,SE_{2t+1})} - \epsilon^{(PE_t,SE_{21t+1})} = (1+\beta)log(\theta w_t + a_t + b_t(1-r_b)) + log\frac{r}{r+\alpha} + \beta log\frac{r^\alpha\theta\alpha}{r+\alpha}$$

- $\{SE_{2t}, SE_{2t+1}\}$

$$U^{(SE_{2t},SE_{2t+1})} - \epsilon^{(SE_{2t},SE_{2t+1})} = (1+\beta)log(a_t^\alpha) + log\frac{r\theta}{r+\alpha} + \beta log\frac{r^{\alpha+1}\theta^2}{r+\alpha}$$

- $\{SE_{2t}, SE_{1t+1}\}$

$$U^{(SE_{2t},SE_{1t+1})} - \epsilon^{(SE_{2t},SE_{1t+1})} = (1+\beta)log\left(\frac{r\theta(a_t)^\alpha + \theta\gamma^\alpha - r\gamma - \pi_{t+1}^{SE_1} - \psi_{t+1}}{(r^2+r)}\right) + \beta log(r)$$

- $\{SE_{2t}, PE_{t+1}\}$

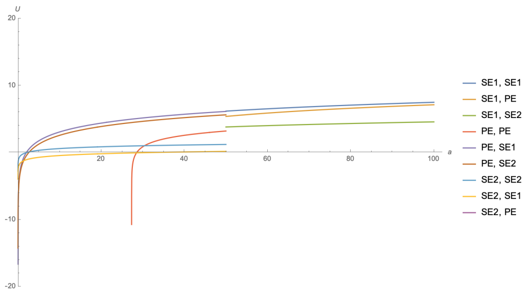
$$U^{(SE_{2t},PE_{t+1})} - \epsilon^{(SE_{2t},PE_{t+1})} = (1+\beta)log\left(\frac{r\theta(a_t)^\alpha + \theta w_{t+1} + b_{t+1}(1-r_b) - \psi_{t+1}}{(r+1)}\right)$$

[22] If $\pi_{t+1}^{PE} \neq 0$ it's impossible to find implicit functional form for $U^{(PE_t,SE_{2t+1})}$.

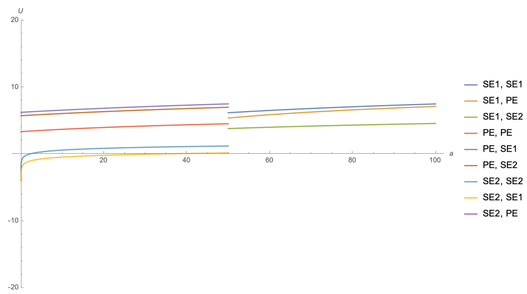
Relative Positions of Utility Functions

Let's consider an economy where $\beta = 0.9$, $\theta = 1.03$[23], $r = 1 + 0.025$[24], and assume that the production function θk^{α} has an decreasing returns to scale, $\alpha = 0.2$. In this economy $\gamma = 1.04$.

Fig. (a) shows the initial position of all utilities function ($a_b = 50$, $\gamma = 1.04$, $\alpha = 0.2$, $\theta = 1.03$, $r = 1 + 0.025 = 1.025$, $\beta = 0.9$, $r_b = 1.04$, $\psi_t = 0$, $\psi_{t+1} = 0$, $w_t = 0$, $w_{t+1} = 0$, $\pi_{t+1}^{PE} = 0.5$, $\pi_{t+1}^{SE1} = 0.5$). Under these assumptions about coefficients, the agent chooses $\{PE_t, SE_{t+1}\}$ if $a_t \leq a_b$ and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$.



a)



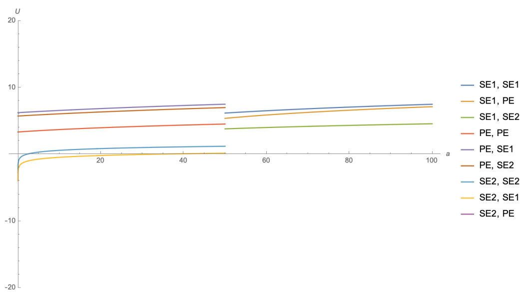
b)

Fig. (b) and (c) show positions of all utilities function if wages (for paid-employed worker) increase keeping all other parameters at the minimal level. Fig. (b) $w_t \uparrow$, Fig.(c) $w_{t+1} \uparrow$. ($a_b = 50$, $\gamma = 1.04$, $\alpha = 0.2$, $\theta = 1.03$, $r = 1 + 0.025 = 1.025$, $\beta = 0.9$, $r_b = 1.04$, $\psi_t = 0$, $\psi_{t+1} = 0$, $w_t \geq 0$ (Fig.(b)), $w_{t+1} \geq 0$ (Fig.(c)), $\pi_{t+1}^{PE} = 0.5$, $\pi_{t+1}^{SE1} = 0.5$). If wages increase in the first period, $w_t \uparrow$, the agent chooses $\{PE_t, SE_{1t+1}\}$ if $a_t \leq a_b$, and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$. Fig.(b) also shows such increase in $w_t \uparrow$ that the position of $U^{(PE_t, SE_{1t+1})}$ is above $U^{(SE_{1t}, SE_{1t+1})}$. It's possible that the agent prefers to be paid-employed worker in period t and Type-1 self-employed in period $t+1$ regardless the amount of initial wealth, a_t. If wages increase in the second period, $w_{t+1} \uparrow$, the agent chooses $\{SE_{2t}, PE_{t+1}\}$ if $a_t \leq a_b$, and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$. Fig.(c) shows the

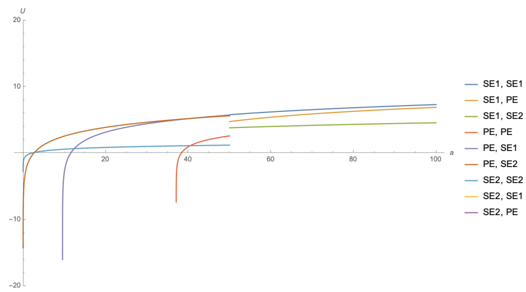
[23] University of Groningen and University of California, Davis, Total Factor Productivity at Constant National Prices for United States, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/RTFPNAUSA632NRUG, August 8, 2019.

[24] The Federal Reserve Interest Rate Decision on June, 2019

situation when the agent prefers to work as Type-2 self-employed at period t and paid-employed worker at period $t+1$ regardless the amount of initial wealth, a_t.

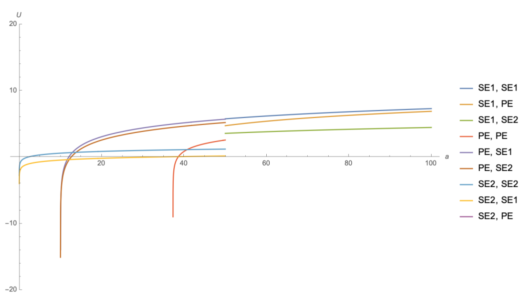


c)

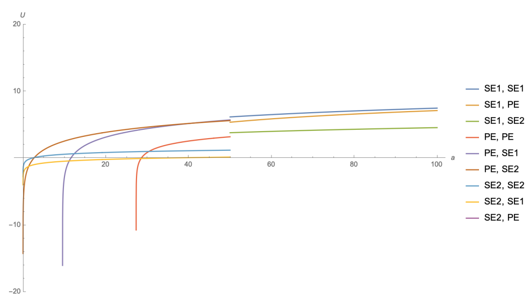


d)

Fig. (d) and (e) show positions of all utilities function if the non-interest cost of borrowing increases keeping all other parameters at the minimal level. Fig. (d) $\psi_t \uparrow$, Fig.(e) $\psi_{t+1} \uparrow$. ($a_b = 50$, $\gamma = 1.04$, $\alpha = 0.2$, $\theta = 1.03$, $r = 1 + 0.025 = 1.025$, $\beta = 0.9$, $r_b = 1.04$, $\psi_t \uparrow$ (Fig.(d)), $\psi_{t+1} \uparrow$ (Fig.(e)), $w_t = 0$, $w_{t+1} = 0$, $\pi^{PE}_{t+1} = 0$, $\pi^{SE1}_{t+1} = 0$). If the non-interest cost of borrowing increases in the first period, $\psi_t \uparrow$, the agent chooses $\{SE_{2t}, SE_{2t+1}\}$ if $a_t \leq a_1 \leq a_b$, $\{PE_t, SE_{2t+1}\}$ if $a_1 \leq a_t \leq a_2 \leq a_b$, $\{PE_t, SE_{1t+1}\}$ if $a_2 \leq a_b$ (where a_1 and a_2 are intersections of utility functions), and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$. If the non-interest cost of borrowing increases in the second period, $\psi_{t+1} \uparrow$, the agent chooses $\{SE_{2t}, SE_{2t+1}\}$ if $a_t \leq a_1 \leq a_b$, $\{PE_t, SE_{1t+1}\}$ if $a_1 \leq a_t \leq a_b$, and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$.



e)



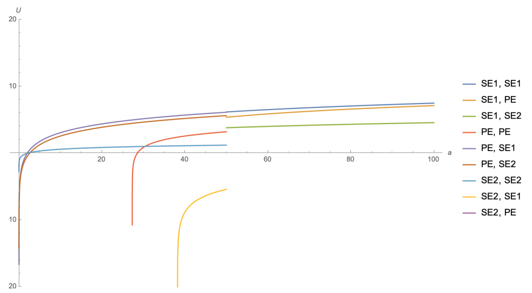
f)

Fig. (f) shows positions of all utilities function if the cost of switching from paid-employment, π^{PE}_{t+1} increases keeping all other parameters at the minimal level. ($a_b = 50$,

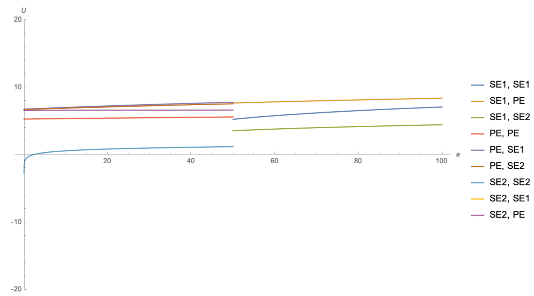
$\gamma = 1.04$, $\alpha = 0.2$, $\theta = 1.03$, $r = 1 + 0.025 = 1.025$, $\beta = 0.9$, $r_b = 1.04$, $\psi_t = 0$, $\psi_{t+1} = 0$, $w_t = 0$, $w_{t+1} = 0$, $\pi^{PE}_{t+1} \uparrow$, $\pi^{SE1}_{t+1} = 0$). If the cost of switching from paid-employment increases, $\pi^{PE}_{t+1} \uparrow$, the agent chooses $\{SE_{2t}, SE_{2t+1}\}$ if $a_t \leq a_1 \leq a_b$, $\{PE_t, SE_{2t+1}\}$ if $a_1 \leq a_t \leq a_b$, and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$.

Fig. (g) shows positions of all utilities function if the cost of switching to Type-1 self-employment, π^{SE1}_{t+1}, increases keeping all other parameters at the minimal level. ($a_b = 50$, $\gamma = 1.04$, $\alpha = 0.2$, $\theta = 1.03$, $r = 1 + 0.025 = 1.025$, $\beta = 0.9$, $r_b = 1.04$, $\psi_t = 0$, $\psi_{t+1} = 0$, $w_t = 0$, $w_{t+1} = 0$, $\pi^{PE}_{t+1} = 0$, $\pi^{SE1}_{t+1} \uparrow$). If the cost of switching to Type-1 self-employment increases, $\pi^{SE1}_{t+1} \uparrow$, the agent chooses $\{SE_{2t}, SE_{2t+1}\}$ if $a_t \leq a_1 \leq a_b$, $\{PE_t, SE_{1t+1}\}$ if $a_1 \leq a_t \leq a_b$, and $\{SE_{1t}, SE_{1t+1}\}$ if $a_t > a_b$.

Fig. (h) shows positions of all utilities function if all costs increase ($a_b = 50$, $\gamma = 1.04$, $\alpha = 0.2$, $\theta = 1.03$, $r = 1 + 0.025 = 1.025$, $\beta = 0.9$, $r_b = 1.04$, $\psi_t \uparrow$, $\psi_{t+1} \uparrow$, $w_t \uparrow$, $w_{t+1} \uparrow$, $\pi^{PE}_{t+1} \uparrow$, $\pi^{SE1}_{t+1} \uparrow$). In this case the agent chooses $\{PE_t, SE_{2t+1}\}$ if $a_t \leq a_1 \leq a_b$, $\{PE_t, SE_{1t+1}\}$ if $a_1 \leq a_t \leq a_b$, and $\{SE_{1t}, PE_{t+1}\}$ if $a_t > a_b$.



g)



h)

The Role of Non-Interest Cost of Borrowing ψ_t in the Decision of Being Type-1 Self-Employed at Period t

The agent prefers to being as Type-1 self-employed at period t and $t+1$ if she receives the highest utility function:

$$U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(PE_t,SE_{1t+1})}, U^{(SE_{2t},SE_{1t+1})}, U^{(PE_t,SE_{2t+1})}, U^{(SE_{2t},SE_{2t+1})}, U^{(PE_t,PE_{t+1})}, U^{(SE_{2t},SE_{2t+1})}, U^{(SE_{1t},PE_{t+1})}, U^{(SE_{1t},SE_{2t+1})}\}$$

To investigate the role of non-interest cost of borrowing ψ_t in the decision of being Type-1 self-employed at period t, I fix a labor choice at the period $(t+1)$ as Type-1 self-employed for simplicity to avoid an estimation of all nine possible labor choices. The extension of the model for other labor choices PE_{t+1} and SE_{2t+1} is straightforward.

The agent prefers to works as Type-1 self-employed at period t if she receives the highest utility function:

$$U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(PE_t,SE_{1t+1})}, U^{(SE_{2t},SE_{1t+1})}\}$$

In this subsection I show that the probability of being Type-1 self-employed at period t declines if non-interest cost of borrowing at period t increases.

$$\frac{\partial Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(PE_t,SE_{1t+1})}, U^{(SE_{2t},SE_{1t+1})}\}\right)}{\partial \psi_t} \leq 0$$

The agent's task if she works as Type-1 self-employed at period t and $t+1$:

$$\max_{c_t \geq 0, c_{t+1} \geq 0} \mathbb{E}U^{(SE_{1t},SE_{1t+1})} = U(c_t) + \beta U(c_{t+1})$$

Agent's budget constraints ($a_t \geq a_b$ and $a_{t+1} \geq a_b$):

$$\text{Period } t: \quad a_{t+1} + c_t = \theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t \tag{9}$$

$$\text{Period } t+1: \quad c_{t+1} = \theta\gamma^{\alpha} - r(\gamma - ra_{t+1}) - \psi_{t+1} \tag{10}$$

where $\gamma = \left(\frac{r}{\lambda\theta}\right)^{\frac{1}{\alpha-1}}$, and $a_t \geq a_b$.

The first order condition for an internal solution of the agent's problem leads to the following Euler equation:

$$U^{'}(c_t) = \beta r^2 U^{'}(c_{t+1})$$

Assume that $\beta r = 1$ and $U(c_t) = log(c_t)$.

$$rc_t = c_{t+1} \tag{11}$$

Substitute the eq.11 in budget constraints eq.10 and eq.9:

$$r(\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - a_{t+1}) = \theta\gamma^{\alpha} - r(\gamma - ra_{t+1}) - \psi_{t+1} \Rightarrow$$
$$a_{t+1} = \frac{\theta\gamma^{\alpha}(r-1) + r\gamma + \psi_{t+1} - r^2(\gamma - a_t) - r\psi_t}{r^2 + r}$$

The agent's total utility function:

$$U^{(SE_{1t},SE_{1t+1})} - \epsilon^{(SE_{1t},SE_{1t+1})} = log(c_t) + \beta log(c_{t+1}) = log(c_t)(1+\beta) + \beta log(r) =$$
$$= log(\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - \frac{\theta\gamma^{\alpha}(r-1) + r\gamma + \psi_{t+1} - r^2(\gamma - a_t) - r\psi_t}{r^2 + r})(1+\beta) + \beta log(r) =$$
$$= log(A - \frac{\psi_{t+1}}{r^2+r} - \frac{r^2\psi_t}{r^2+r})(1+\beta) + \beta log(r) \Rightarrow$$
$$\frac{\partial U^{(SE_{1t},SE_{1t+1})}}{\partial \psi_t} = \frac{(1+\beta)\frac{-r}{r+1}}{A - \frac{\psi_{t+1}}{r^2+r} - \frac{r^2\psi_t}{r^2+r}} \leq 0, \quad \frac{\partial U^{(SE_{1t},SE_{1t+1})}}{\partial \psi_{t+1}} = \frac{(1+\beta)\frac{-1}{r^2+r}}{A - \frac{\psi_{t+1}}{r^2+r} - \frac{r^2\psi_t}{r^2+r}} \leq 0$$

where $A = \theta\gamma^{\alpha} - r(\gamma - a_t) - \frac{\theta\gamma^{\alpha}(r-1) + r\gamma - r^2(\gamma - a_t)}{r^2+r} = \frac{r^2(\theta\gamma^{\alpha} - r\gamma) + r^2 a_t(r-1) - (\theta\gamma^{\alpha} + r\gamma)}{r^2+r}$.

The increasing in the non-interest cost of borrowing in the period t or period $t+1$ decreases the total utility function.

The agent's task if she works as paid-employed at period t and as Type-1 self-employed worker at period $t+1$:

$$\max_{c_t \geq 0, c_{t+1} \geq 0} \mathbb{E}U^{(PE_t,SE_{1t+1})} = U(c_t) + \beta U(c_{t+1})$$

Agent's budget constraints ($a_t \leq a_b$ and $a_{t+1} \geq a_b$):

$$\text{Period } t: \quad a_{t+1} + c_t = \theta w_t + a_t + b_t(1 - r_b) - \psi_t \tag{12}$$

$$\text{Period } t+1: \quad c_{t+1} = \theta\gamma^{\alpha} - r\left(\gamma - ra_{t+1}\right) - \psi_{t+1} - \pi^{PE}_{t+1} \tag{13}$$

The first order condition for an internal solution of the agent's problem leads to Euler equation:

$$U^{'}(c_t) = \beta r^2 U^{'}(c_{t+1})$$

Assume $\beta r = 1$, and $U(c_t) = log(c_t)$.

$$rc_t = c_{t+1} \tag{14}$$

Substitute the eq.14 in budget constraints eq.13 and eq.12 :

$$r(\theta w_t + b_t(1 - r_b) - \psi_t + a_t - a_{t+1}) = \theta\gamma^{\alpha} - r\left(\gamma - ra_{t+1}\right) - \psi_{t+1} - \pi^{PE}_{t+1} \Rightarrow$$
$$a_{t+1} = \frac{r\theta w_t + rb_t(1 - r_b) - r\psi_t + ra_t - \theta\gamma^{\alpha} + r\gamma + \psi_{t+1} + \pi^{PE}_{t+1}}{(r^2 + r)}$$

The agent's total utility function:

$$U^{(PE_t, SE_{1t+1})} - \epsilon^{(PE_t, SE_{1t+1})} = log(c_t) + \beta log(c_{t+1}) = (1+\beta)log(c_t) + \beta log(r) = (1+\beta)\cdot$$
$$\cdot\, log\left(\theta w_t + b_t(1 - r_b) - \psi_t + a_t - \frac{r\theta w_t + rb_t(1 - r_b) - r\psi_t + ra_t - \theta\gamma^{\alpha} + r\gamma + \psi_{t+1} + \pi^{PE}_{t+1}}{(r^2 + r)}\right)$$
$$+ \beta log(r) =$$
$$= (1+\beta)log\left(\frac{r^2\theta w_t + r^2 a_t + r^2(b_t(1 - r_b) - \psi_t) + \theta\gamma^{\alpha} - r\gamma - \psi_{t+1} - \pi^{PE}_{t+1}}{(r^2 + r)}\right) + \beta log(r) =$$
$$= (1+\beta)log\left(G - \frac{r^2\psi_t}{(r^2 + r)} - \frac{\psi_{t+1}}{(r^2 + r)}\right) + \beta log(r) \Rightarrow$$
$$\frac{\partial U^{(PE_t, SE_{1t+1})}}{\partial \psi_{t+1}} = \frac{(1+\beta)\frac{-r}{r+1}}{G - \frac{\psi_{t+1}}{r^2+r} - \frac{r^2\psi_t}{r^2+r}} \leq 0; \quad \frac{\partial U^{(PE_t, SE_{1t+1})}}{\partial \psi_t} = \frac{(1+\beta)\frac{-1}{r^2+r}}{G - \frac{\psi_{t+1}}{r^2+r} - \frac{r^2\psi_t}{r^2+r}} \leq 0$$

where $G = \frac{r^2\theta w_t + r^2 a_t + r^2 b_t(1 - r_b) + \theta\gamma^{\alpha} - r\gamma - \pi^{PE}_{t+1}}{(r^2+r)}$.

The agent's task if she worked as Type-2 self-employed at period t and as Type-1 self-employed worker at period $t+1$:

$$\max_{c_t \geq 0, c_{t+1} \geq 0} U(c_t) + \beta U(c_{t+1})$$

Agent's budget constraints ($a_t \leq a_b$ and $a_{t+1} \geq a_b$):

$$\text{Period } t: \quad \text{s.t } a_{t+1} + c_t = \theta(a_t)^{\alpha} \tag{15}$$

$$\text{Period } t+1: \quad c_{t+1} = \theta\gamma^{\alpha} - r\left(\gamma - ra_{t+1}\right) - \psi_{t+1} - \pi_{t+1}^{SE_1} \tag{16}$$

The first order condition for an internal solution of the agent's problem leads to Euler equation:

$$U^{'}(c_t) = \beta r^2 U^{'}(c_{t+1})$$

Under the assumption $\beta r = 1$, and $U(c_t) = log(c_t)$.

$$rc_t = c_{t+1} \tag{17}$$

Substitute the eq.17 in budget constraints eq.15 and eq.16:

$$\begin{aligned} & r(\theta(a_t)^{\alpha} - a_{t+1}) = \theta\gamma^{\alpha} - r\left(\gamma - ra_{t+1}\right) - \psi_{t+1} - \pi_{t+1}^{SE_1} \Rightarrow \\ & a_{t+1} = \frac{r\theta(a_t)^{\alpha} - \theta\gamma^{\alpha} + r\gamma + \psi_{t+1} + \pi_{t+1}^{SE_1}}{(r^2+r)} \end{aligned}$$

The agent's total utility function:

$$\begin{aligned} & U^{(SE_{2t}, SE_{1t+1})} - \epsilon^{(SE_{2t}, SE_{1t+1})} = log(c_t) + \beta log(c_{t+1}) = (1+\beta) log(c_t) + \beta log(r) = \\ & = (1+\beta) log\left(\theta(a_t)^{\alpha} - \frac{r\theta(a_t)^{\alpha} - \theta\gamma^{\alpha} + r\gamma + \psi_{t+1} + \pi_{t+1}^{SE_1}}{(r^2+r)}\right) + \beta log(r) = \\ & = (1+\beta) log\left(\frac{r\theta(a_t)^{\alpha} + \theta\gamma^{\alpha} - r\gamma - \psi_{t+1} - \pi_{t+1}^{SE_1}}{(r^2+r)}\right) + \beta log(r) = \\ & = (1+\beta) log\left(H - \frac{\psi_{t+1}}{(r^2+r)}\right) + \beta log(r) \end{aligned}$$

where $H = \frac{r\theta(a_t)^{\alpha} + \theta\gamma^{\alpha} - r\gamma - \pi_{t+1}^{SE_1}}{(r^2+r)}$.

Assume that $\epsilon^j \in N(0;1)$, the probability of Type-1 self-employment at period t:

$$Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(PE_t,SE_{1t+1})}, U^{(SE_{2t},SE_{1t+1})}\}\right) =$$

$$= Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq U^{(PE_t,SE_{1t+1})}\right) \cdot Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq U^{(SE_{2t},SE_{1t+1})}\right) =$$

$$= Pr((1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) + \beta log(r) + \epsilon^{(SE_{1t},SE_{1t+1})} \geq$$

$$\geq (1+\beta)log\left(G - \frac{r^2\psi_t}{(r^2+r)} - \frac{\psi_{t+1}}{(r^2+r)}\right) + \beta log(r) + \epsilon^{(PE_t,SE_{1t+1})})\cdot$$

$$\cdot Pr((1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) + \beta log(r) + \epsilon^{(SE_{1t},SE_{1t+1})} \geq$$

$$\geq (1+\beta)log\left(H - \frac{\psi_{t+1}}{(r^2+r)}\right) + \beta log(r) + \epsilon^{(PE_t,SE_{1t+1})}) =$$

$$= \Phi\left(\frac{1}{\sqrt{2}}\left(\underbrace{(1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) - (1+\beta)log\left(G - \frac{r^2\psi_t}{(r^2+r)} - \frac{\psi_{t+1}}{(r^2+r)}\right)}_{=J}\right)\right)\cdot$$

$$\cdot \Phi\left(\frac{1}{\sqrt{2}}\left(\underbrace{(1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) - (1+\beta)log\left(H - \frac{\psi_{t+1}}{(r^2+r)}\right)}_{=K}\right)\right)$$

$$\frac{\partial Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(PE_t,SE_{1t+1})}, U^{(SE_{2t},SE_{1t+1})}\}\right)}{\partial \psi_t} = \Phi^{'}(J)\cdot(1+\beta)\frac{1}{\sqrt{2}}(\frac{\left(-\frac{r}{r+1}\right)}{A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}} -$$

$$- \frac{-\frac{r}{r+1}}{G - \frac{r^2\psi_t}{(r^2+r)} - \frac{\psi_{t+1}}{(r^2+r)}}) \cdot \Phi(K) + \Phi^{'}(K)\cdot(1+\beta)\frac{1}{\sqrt{2}}\frac{\left(-\frac{r}{r+1}\right)}{A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}}\cdot\Phi(J) =$$

$$= \underbrace{\Phi^{'}(J)\cdot(1+\beta)\frac{1}{\sqrt{2}}\left(-\frac{r}{r+1}\right)\frac{G-A}{(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)})(G - \frac{r^2\psi_t}{(r^2+r)} - \frac{\psi_{t+1}}{(r^2+r)})}\Phi(K) +}_{\leq ? \geq 0}$$

$$\underbrace{+ \Phi^{'}(K)\cdot(1+\beta)\frac{1}{\sqrt{2}}\frac{\left(-\frac{r}{r+1}\right)}{A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}}\cdot\Phi(J)}_{\leq 0}$$

Let's consider $(G-A)$ separately:

$$G - A = \frac{r^2\theta w_t + r^2 a_t + r^2 b_t(1-r_b) + \theta\gamma^\alpha - r\gamma - \pi^{PE}_{t+1}}{(r^2+r)} - \frac{r^2(\theta\gamma^\alpha - r\gamma) + r^2 a_t(r-1) - (\theta\gamma^\alpha + r\gamma)}{r^2+r} =$$

$$= \frac{r^2\theta w_t + a_t r^2(2-r) + r^2 b_t(1-r_b) + \theta\gamma^\alpha(2-r^2) + r^2\gamma - \pi^{PE}_{t+1}}{r^2+r}$$

It can be assumed that $r \leq 1.4 \leq \sqrt{2}$, because r is one plus the interest rate, and probably the interest rate will not exceed 40%. Also it can be assumed that $\pi_{t+1}^{PE} \leq r^2\theta w_t + a_t r^2(2-r) + r^2 b_t(1-r_b) + \theta\gamma^\alpha(2-r^2) + r^2\gamma$. So $G - A \geq 0 \Rightarrow G \geq A$.

If $G \geq A$, then $|\frac{\partial U^{(SE_{1t}, SE_{1t+1})}}{\partial \psi_{t+1}}| \geq |\frac{\partial U^{(PE_t, SE_{1t+1})}}{\partial \psi_{t+1}}|$, and $\frac{\partial Pr\left(U^{(SE_{1t}, SE_{1t+1})} \geq \max\{U^{(PE_t, SE_{1t+1})}, U^{(SE_{2t}, SE_{1t+1})}\}\right)}{\partial \psi_t} \leq 0$.

The probability of being Type-1 self-employment at period t declines if non-interest cost of borrowing ψ_t increases.

The Role of Non-Interest Cost of Borrowing ψ_{t+1} in the Decision of Being Type-1 Self-Employed at Period $t+1$

To investigate the role of non-interest cost of borrowing ψ_{t+1} in the decision of being Type-1 self-employed at period $t+1$, I fix a labor choice at the period t as Type-1 self-employed for simplicity to avoid an estimation of all nine possible labor choices.. The extension of the model for other labor choices PE_t and SE_{2t} is straightforward.

The agent prefers to works as Type-1 self-employed at period $t+1$ if she receives the highest utility function:

$$U^{(SE_{1t}, SE_{1t+1})} \geq \max\{U^{(SE_{1t}, PE_{t+1})}, U^{(SE_{1t}, SE_{2t+1})}\}$$

In this subsection I show that the probability of being Type-1 self-employed at period $t+1$ declines if non-interest cost of borrowing at period $t+1$ increases.

$$\frac{\partial Pr\left(U^{(SE_{1t}, SE_{1t+1})} \geq \max\{U^{(SE_{1t}, PE_{t+1})}, U^{(SE_{1t}, SE_{2t+1})}\}\right)}{\partial \psi_{t+1}} \leq 0$$

The total utility of being Type-1 self-employed at period t and $t+1$ is described in the previous subsection.

The agent's task if she works as Type-1 self-employed at period t and as paid-employed worker at period $t+1$:

$$\max_{c_t \geq 0, c_{t+1} \geq 0} U(c_t) + \beta U(c_{t+1})$$

Agent's budget constraints ($a_t \geq a_b$ and $a_{t+1} \leq a_b$):

$$\text{Period } t: \quad a_{t+1} + c_t = \theta\gamma^{\alpha} - r\left(\gamma - a_t\right) - \psi_t \tag{18}$$

$$\text{Period } t+1: \quad c_{t+1} = \theta w_{t+1} + r a_{t+1} + b_{t+1}(1 - r_b) - \psi_{t+1} \tag{19}$$

The first order condition for an internal solution of the agent's problem leads to Euler equation:

$$U^{'}(c_t) = \beta r U^{'}(c_{t+1})$$

Under the assumption $\beta r = 1$, and $U(c_t) = log(c_t)$.

$$c_t = c_{t+1} \tag{20}$$

Substitute the eq.20 in budget constraints eq.18 and eq.19:

$$\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - a_{t+1} = \theta w_{t+1} + r a_{t+1} + b_{t+1}(1 - r_b) - \psi_{t+1} \Rightarrow$$
$$a_{t+1} = \frac{\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - \theta w_{t+1} - b_{t+1}(1 - r_b) + \psi_{t+1}}{(r+1)}$$

The agent's total utility function:

$$U^{(SE_{1t}, PE_{t+1})} - \epsilon^{(SE_{1t}, PE_{t+1})} = log(c_t) + \beta log(c_{t+1}) = (1+\beta) log(c_t) =$$
$$= (1+\beta) log\left(\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - \frac{\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - \theta w_{t+1} - b_{t+1}(1 - r_b) + \psi_{t+1}}{(r+1)}\right) =$$
$$= (1+\beta) log(L - \frac{\psi_{t+1}}{r+1} - \frac{r\psi_t}{r+1}) \Rightarrow \frac{\partial U^{(SE_{1t}, PE_{t+1})}}{\partial \psi_t} \leq 0; \ \frac{\partial U^{(SE_{1t}, PE_{t+1})}}{\partial \psi_{t+1}} \leq 0$$

where $L = \theta\gamma^{\alpha} - r(\gamma - a_t) - \frac{\theta\gamma^{\alpha} - r(\gamma - a_t) - \theta w_{t+1} - b_{t+1}(1 - r_b)}{(r+1)} = \frac{r(\theta\gamma^{\alpha} - r\gamma) + r^2 a_t + \theta w_{t+1} + b_{t+1}(1 - r_b)}{(r+1)} = \frac{r^2(\theta\gamma^{\alpha} - r\gamma) + r^3 a_t + r\theta w_{t+1} + r b_{t+1}(1 - r_b)}{(r^2+r)} = A - \frac{r^3 a_t - (\theta\gamma^{\alpha} - r\gamma) - r\theta w_{t+1} - r b_{t+1}(1 - r_b)}{(r^2+r)}$

The agent's task if she works as Type-1 self-employed at period t and as Type-2 self-employed worker at period $t+1$:

$$\max_{c_t \geq 0, c_{t+1} \geq 0} U(c_t) + \beta U(c_{t+1})$$

Agent's budget constraints ($a_t \geq a_b$ and $a_{t+1} \leq a_b$):

$$\text{Period } t: \quad a_{t+1} + c_t = \theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t \tag{21}$$

$$\text{Period } t+1: \quad c_{t+1} = \theta(r a_{t+1})^{\alpha} \tag{22}$$

The first order condition for an internal solution of the consumer's problem leads to Euler equation:

$$U^{'}(c_t) = \beta r U^{'}(c_{t+1})(r^{\alpha-1}\theta\alpha(\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t - c_t)^{\alpha-1})$$

Under the assumption $\beta r = 1$, $M = [\theta\gamma^{\alpha} - r(\gamma - a_t) - \psi_t]$ and $U(c_t) = log(c_t)$.

$$r^{\alpha-1}\theta_2\alpha(M - c_t)^{\alpha-1}c_t = c_{t+1} \tag{23}$$

Substitute the eq.23 in budget constraints eq.21 and eq.22:

$$\theta(ra_{t+1})^{\alpha} = r^{\alpha-1}\theta\alpha(a_{t+1})^{\alpha-1}(M - a_{t+1}) \Rightarrow a_{t+1} = \frac{\alpha M}{\alpha + r}$$

The agent's total utility function:

$$U^{(SE_{1t},SE_{2t+1})} - \epsilon^{(SE_{1t},SE_{2t+1})} = log(c_t) + \beta log(c_{t+1}) = (1+\beta)log(c_t) + log(r^{\alpha-1}\theta\alpha(M - c_t)^{\alpha-1}) =$$
$$= \underbrace{(1+\beta)log\frac{rM}{r+\alpha} + (\alpha-1)log\frac{\alpha M}{r+\alpha} + log(r^{\alpha-1}\theta\alpha)}_{=N\geq 0} \Rightarrow \frac{\partial U^{(SE_{1t},SE_{2t+1})}}{\partial\psi_t} \leq 0$$

Assume that $\epsilon^j \in N(0;1)$, the probability of being Type-1 self-employment at period $t+1$:

$$Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(SE_{1t},PE_{t+1})}, U^{(SE_{1t},SE_{2t+1})}\}\right) =$$
$$Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq U^{(SE_{1t},PE_{t+1})}\right) \cdot Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq U^{(SE_{1t},SE_{2t+1})}\right) =$$
$$= Pr((1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) + \beta log(r) + \epsilon^{(SE_{1t},SE_{1t+1})} \geq$$
$$\geq (1+\beta)log\left(L - \frac{\psi_{t+1}}{r+1} - \frac{r\psi_t}{r+1}\right) + \epsilon^{(SE_{1t},PE_{t+1})}).$$
$$\cdot Pr((1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) + \beta log(r) + \epsilon^{(SE_{1t},SE_{1t+1})} \geq$$
$$\geq N + \epsilon^{(SE_{1t},SE_{2t+1})}) =$$
$$= \Phi\left(\frac{1}{\sqrt{2}}\left(\underbrace{(1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) + \beta log(r) - (1+\beta)log\left(L - \frac{\psi_{t+1}}{r+1} - \frac{r\psi_t}{r+1}\right)}_{=O}\right)\right) \cdot$$
$$\cdot \Phi\left(\frac{1}{\sqrt{2}}\left(\underbrace{(1+\beta)log\left(A - \frac{\psi_{t+1}}{(r^2+r)} - \frac{r^2\psi_t}{(r^2+r)}\right) + \beta log(r) - N}_{=P}\right)\right)$$

$$\frac{\partial Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(SE_{1t},PE_{t+1})},U^{(SE_{1t},SE_{2t+1})}\}\right)}{\partial \psi_{t+1}} = \Phi^{'}(O)\cdot(1+\beta)\frac{1}{\sqrt{2}}(\frac{\left(-\frac{1}{r^2+r}\right)}{A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)}}-$$

$$-\frac{\left(-\frac{1}{r+1}\right)}{L-\frac{\psi_{t+1}}{r+1}-\frac{r\psi_t}{r+1}})\cdot\Phi(P)+\Phi^{'}(P)\cdot(1+\beta)\frac{1}{\sqrt{2}}\frac{\left(-\frac{1}{r^2+1}\right)}{A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)}}\cdot\Phi(O)=$$

$$=\Phi^{'}(O)\cdot(1+\beta)\frac{1}{\sqrt{2}}\left(\frac{-1}{r^2+r}\right)\left(\frac{(L-\frac{\psi_{t+1}}{r+1}-\frac{r\psi_t}{r+1})-r(A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)})}{(A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)})(L-\frac{\psi_{t+1}}{r+1}-\frac{r\psi_t}{r+1})}\right)\cdot\Phi(P)+$$

$$+\Phi^{'}(P)\cdot(1+\beta)\frac{1}{\sqrt{2}}\frac{\left(-\frac{1}{r^2+1}\right)}{A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)}}\cdot\Phi(O)=$$

$$=\underbrace{\Phi^{'}(O)\cdot(1+\beta)\frac{1}{\sqrt{2}}(\frac{L-rA}{(A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)})(L-\frac{\psi_{t+1}}{r+1}-\frac{r\psi_t}{r+1})})\cdot\Phi(P)+}_{\leq ?\geq 0}$$

$$\underbrace{+\Phi^{'}(P)\cdot(1+\beta)\frac{1}{\sqrt{2}}\frac{\left(-\frac{1}{r^2+1}\right)}{A-\frac{\psi_{t+1}}{(r^2+r)}-\frac{r^2\psi_t}{(r^2+r)}}\cdot\Phi(O)}_{\leq 0}$$

Let's consider $(L-rA)$ separately:

$$L-rA=\frac{r^2(\theta\gamma^\alpha-r\gamma)+r^3a_t+r\theta w_{t+1}+rb_{t+1}(1-r_b)}{(r^2+r)}-$$
$$-r\frac{r^2(\theta\gamma^\alpha-r\gamma)+r^2a_t(r-1)-(\theta\gamma^\alpha+r\gamma)}{r^2+r}=$$
$$=\frac{r^2(\theta\gamma^\alpha-r\gamma)(r-1)+r^4a_t+r\theta w_{t+1}+rb_{t+1}(1-r_b)}{(r^2+r)}\geq 0$$

It means:

$$\frac{\partial Pr\left(U^{(SE_{1t},SE_{1t+1})} \geq \max\{U^{(SE_{1t},PE_{t+1})},U^{(SE_{1t},SE_{2t+1})}\}\right)}{\partial \psi_{t+1}}\leq 0$$

The probability of being Type-1 self-employment at period $t+1$ declines if non-interest cost of borrowing ψ_{t+1} increases.

The Role of Non-Interest Cost of Borrowing ψ_{t+1} in the Decision of Being Paid-Employed at Period $t+1$

In this subsection I want to show that if the agent is Type-2 self-employed worker (which means she does not have access to credit market), she may prefer to switch to paid-employment to be able to use credit products (e.g., mortgage) if non-interest cost of borrowing declines. To show it I complicate the model by allowing for paid-employed to take a one period loan in the bank. Paid-employed worker budget constraints:

$$\text{Period } t\text{:} \quad a_{t+1} + c_t = \theta w_t + a_t + b_t(1 - r_b) - \psi_t$$

$$\text{Period } t+1\text{:} \quad c_{t+1} = \theta w_{t+1} + r a_{t+1} + b_{t+1}(1 - r_b) - \psi_{t+1}$$

To investigate the role of non-interest cost of borrowing ψ_{t+1} in the decision of Type-2 self-employed at period t to switch to paid-employment at period $t+1$ - $SE_{2t} \to PE_{t+1}$, I consider all possible labor choices that the agent has at period $t+1$ if she was Type-2 self-employed at period t -it's SE_{1t+1} and SE_{2t+1}:

$$Pr(SE_{2t} \to PE_{t+1}) = Pr(U^{(SE_{2t},PE_{t+1})} \geq \max\{U^{(SE_{2t},SE_{2t+1})}, U^{(SE_{2t},SE_{1t+1})}\}$$

In this subsection I show that the probability of switching $SE_{2t} \to PE_{t+1}$ declines if non-interest cost of borrowing at period $t+1$ increases.

$$\frac{\partial Pr(SE_{2t} \to PE_{t+1})}{\partial \psi_{t+1}} \leq 0$$

The agent's utility function if she works as Type-2 self-employed worker at period t and paid-employed worker at period $t+1$:

$$U^{(SE_{2t},PE_{t+1})} - \epsilon^{(SE_{2t},PE_{t+1})} = (1+\beta)log\left(\frac{r\theta(a_t)^\alpha + \theta w_{t+1} + b_{t+1}(1-r_b) - \psi_{t+1}}{(r+1)}\right) =$$

$$= (1+\beta)log\left(Q - \frac{\psi_{t+1}}{(r+1)}\right)$$

where $Q = \frac{r\theta(a_t)^\alpha + \theta w_{t+1} + b_{t+1}(1-r_b)}{(r+1)}$

The agent's utility function if she works as Type-2 self-employed worker at period t and $t+1$:

$$U^{(SE_{2t},SE_{2t+1})} - \epsilon^{(SE_{2t},SE_{2t+1})} = (1+\beta)log(a_t^\alpha) + log\frac{r\theta}{r+\alpha} + \beta log\frac{r^{\alpha+1}\theta^2}{r+\alpha} = R$$

where $R = (1+\beta)log(a_t^{\alpha}) + log\frac{r\theta}{r+\alpha} + \beta log\frac{r^{\alpha+1}\theta^2}{r+\alpha}$

The agent's utility function if she works as Type-2 self-employed worker at period t and Type-1 self-employed worker at period $t+1$:

$$U^{(SE_{2t},SE_{1t+1})} - \epsilon^{(SE_{2t},SE_{1t+1})} = (1+\beta)log\left(\frac{r\theta(a_t)^{\alpha} + \theta\gamma^{\alpha} - r\gamma - \pi_{t+1}^{SE_1} - \psi_{t+1}}{(r^2+r)}\right) + log(r) =$$

$$= (1+\beta)log\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right) + log(r)$$

where $S = \frac{r\theta(a_t)^{\alpha} + \theta\gamma^{\alpha} - r\gamma - \pi_{t+1}^{SE_1}}{(r^2+r)}$

Assume that $\epsilon^j \in N(0;1)$, the probability of being Type-2 self-employment at period t and paid-employed worker at period $t+1$:

$$Pr\left(U^{(SE_{2t},PE_{t+1})} \geq \max\{U^{(SE_{2t},SE_{2t+1})}, U^{(SE_{2t},SE_{1t+1})}\}\right) =$$

$$Pr\left(U^{(SE_{2t},PE_{t+1})} \geq U^{(SE_{2t},SE_{2t+1})}\right) \cdot Pr\left(U^{(SE_{2t},PE_{t+1})} \geq U^{(SE_{2t},SE_{1t+1})}\right) =$$

$$= Pr((1+\beta)log\left(Q - \frac{\psi_{t+1}}{(r+1)}\right) + \epsilon^{(SE_{2t},PE_{t+1})} \geq R + \epsilon^{(SE_{2t},SE_{2t+1})}.$$

$$\cdot Pr((1+\beta)log\left(Q - \frac{\psi_{t+1}}{(r+1)}\right) + \epsilon^{(SE_{2t},PE_{t+1})} \geq (1+\beta)log\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right) + log(r) + \epsilon^{(SE_{2t},SE_{1t+1})}) =$$

$$= \Phi\underbrace{\left(\frac{1}{\sqrt{2}}\left((1+\beta)log\left(Q - \frac{\psi_{t+1}}{(r+1)}\right) - R\right)\right)}_{T\geq 0} \cdot$$

$$\cdot \Phi\underbrace{\left(\frac{1}{\sqrt{2}}\left((1+\beta)log\left(Q - \frac{\psi_{t+1}}{(r+1)}\right) - (1+\beta)log\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right) + log(r)\right)\right)}_{U\geq 0}$$

$$\frac{\partial Pr\left(U^{(SE_{2t},PE_{t+1})} \geq \max\{U^{(SE_{2t},SE_{2t+1})}, U^{(SE_{2t},SE_{1t+1})}\}\right)}{\partial\psi_{t+1}} = \underbrace{\Phi'(T) \cdot (1+\beta)\frac{1}{\sqrt{2}}\left(\frac{-\frac{1}{r+1}}{Q - \frac{\psi_{t+1}}{(r+1)}}\right) \cdot \Phi(U)}_{\leq 0} +$$

$$\underbrace{+ \Phi'(U) \cdot (1+\beta)\frac{1}{\sqrt{2}}\left(\frac{-\frac{1}{r+1}}{Q - \frac{\psi_{t+1}}{(r+1)}} - \frac{\left(-\frac{1}{r^2+r}\right)}{S - \frac{\psi_{t+1}}{(r^2+r)}}\right) \cdot \Phi(T)}_{\leq ?\geq 0}$$

Let's consider the last term separately:

$$\left(\frac{-\frac{1}{r+1}}{Q - \frac{\psi_{t+1}}{(r+1)}} - \frac{\left(-\frac{1}{r^2+r}\right)}{S - \frac{\psi_{t+1}}{(r^2+r)}} \right) = \frac{-\frac{1}{r+1} \cdot \left(S - \frac{\psi_{t+1}}{(r^2+r)}\right) + \frac{1}{r^2+r} \cdot \left(Q - \frac{\psi_{t+1}}{(r+1)}\right)}{\left(Q - \frac{\psi_{t+1}}{(r+1)}\right) \left(S - \frac{\psi_{t+1}}{(r^2+r)}\right)} =$$

$$= \frac{r\theta(a_t)^\alpha + \theta w_{t+1} + b_{t+1}(1-r_b) - \psi_{t+1} - r\theta(a_t)^\alpha - \theta\gamma^\alpha + r\gamma + \pi_{t+1}^{SE_1} + \psi_{t+1}}{(r+1)(r^2+r)\left(Q - \frac{\psi_{t+1}}{(r+1)}\right)\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right)} =$$

$$= \frac{\theta w_{t+1} + b_{t+1}(1-r_b) - \theta\gamma^\alpha + r\gamma + \pi_{t+1}^{SE_1}}{(r+1)(r^2+r)\left(Q - \frac{\psi_{t+1}}{(r+1)}\right)\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right)} = \frac{(c_{PE_{t+1}} - ra_{t+1} - \psi_{t+1}) - (c_{SE_{1t+1}} - ra_{t+1} - \psi_{t+1})}{(r+1)(r^2+r)\left(Q - \frac{\psi_{t+1}}{(r+1)}\right)\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right)} =$$

$$\frac{c_{PE_{t+1}} - c_{SE_{1t+1}}}{(r+1)(r^2+r)\left(Q - \frac{\psi_{t+1}}{(r+1)}\right)\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right)} = \frac{exp(U_{PE_{t+1}}) - exp(U_{SE_{1t+1}})}{(r+1)(r^2+r)\left(Q - \frac{\psi_{t+1}}{(r+1)}\right)\left(S - \frac{\psi_{t+1}}{(r^2+r)}\right)} \leq 0$$

The nominator is negative, because the agent receives higher utility being Type-1 self-employed that being paid-employed worker $U_{SE_{1t+1}} \geq U_{PE_{t+1}}$ if all else being equal.

The probability of switching form Type-1 self-employment at period t to paid-employed at period $t+1$ declines if non-interest cost of borrowing ψ_{t+1} increases.

$$\frac{\partial Pr\left(U^{(SE_{2t}, PE_{t+1})} \geq \max\{U^{(SE_{2t}, SE_{2t+1})}, U^{(SE_{2t}, SE_{1t+1})}\}\right)}{\partial \psi_{t+1}} \leq 0$$

Variables Description

Self-employed workers are self-defined category by respondents. Self-employment is their primary job.

Paid-employed workers are self-defined category and comprised of (a) employees in private business, (b) employees in government.

Age, female, year of survey. Self-explanatory.

Household income. It's a gross income, the sum of all HH members income during the last year.

Race. A binary indicator for following categories: White, Black, Hispanic and other (Asian, Native American and etc).

Education. A categorical variable indicating the highest level of schooling completed by a respondent: [1] "Less than High School diploma", [2] "High School diploma", [3] "College graduate ", and [4] "Doctoral or professional degree. The first category is the base category.

Married. = 1 for legally married individuals (including those not living together) and 0 for other categories including single, widowed, divorced, and living together without marriage.

Number of household members. Counts the number of household members who are presently living in the same household.

Number of kids per household. Counts the number of children under the age of 18 currently residing in the same household.

State population. Number of people living in the state, in thousands. Population is taken from the 2010 Census.

Regions. Set of dummies for living in one of the four regions at the time of interview. The regions are Midwest, Northeast, South and West.

First wave fixed effects. Set of dummies for the starting year of the stochastic sequence or the year of entry to the estimation sample.

Distance to the nearest bank, miles. The distance to the nearest bank is the Euclidean distance between the location of the respondent's home and the bank's address. Source:

the FDIC Summary of Deposits (SOD) data, various years.

Average distance to the 10 nearest banks, miles. All distances between banks withing state and the location of respondent's home were sorted, and 10 nearest banks were chosen. The sum of all distances is divided by 10.

Number of banks within 10 miles. Total number of banks withing 10 miles from the location of respondent.

Number of bank offices per 1,000 state population. The number of bank offices includes bank headquarters, credit organizations, branches, supplementary offices, and operational offices, but excludes cash offices, cash desks, and mobile cash units. Source: the FDIC Summary of Deposits (SOD) data, 2010 Census.

Total assets. The amount of non-housing and housing assets adjusted for inflation at the moment of interview.

Average house value. The average house value within zip-code zone of the location of respondent's home. Source: Zillow.

Sample Design

The Community Advantage Panel Survey (CAPS)[25] was funded by the Ford Foundation and overseen by the UNC Center for Community Capital at the University of North Carolina at Chapel Hill (UNC). The survey sample comprises two subsamples (the owners sample and the renters sample), and the survey was designed to collect information about the economic and social experiences of low-to-moderate-income homeowners and renters.

The owners sample comprises a subset of the low-to-moderate-income homeowners whose loans were purchased by Self-Help, a community development financial institution with headquarters in Durham, North Carolina, as part of the Community Advantage Program (CAP). The panel survey for the owners sample was originally planned for a period of six years. After the initial year of owners sample data collection, a matched panel of low-to- moderate-income renters was selected to be interviewed during the remaining five-year period. The owners sample survey was originally planned to include six telephone interviews and two in-home interviews. In 2008, the decision was made to continue the survey beyond its original end date. The Survey Research Unit at UNC conducted the first four years of telephone interviews (2003 to 2006), and RTI International conducted all subsequent telephone data collection activities and three years of in-home interviews (2005, 2008, and 2012) for the owners sample.

For the renters sample survey, a panel of low-to-moderate-income renters was matched to a subset of owners sample members who were living in the same Metropolitan Statistical Areas (MSAs) in 2004. Data collection for the renters sample began one year after the start of data collection for the owners sample and was originally planned as a five-year survey, which was similarly extended in 2008. RTI conducted all in-home and telephone interviews for the renters sample.

Sample construction was carried out over the course of several years between 2001 and 2004. The owners sample is a convenience sample selected from Self-Help's CAP

[25]Full description of the Community Advantage Panel Survey can be found https://communitycapital.unc.edu/files/2017/10/Paper_22929_extendedabstract_1348_0.pdf

loan portfolio at the beginning of the survey period, and the renters sample is a random sample drawn one year later from neighborhoods near those in which urban owners sample members were living at the time of the baseline survey.

Owners sample. A total of 7,223 CAP loans were purchased by Self-Help before or during the baseline CAPS interview period of 2001-2003. These loans were originated during the period of 1999-2003. All of the borrowers whose loans had been purchased were put into calling, screened (efforts were made to exclude retirees and full-time students, and to include only those borrowers who still lived in their CAP properties and retained their CAP mortgages at the time of the baseline survey), and given the opportunity to participate in the baseline survey if they met the screening criteria. A total of 3,743 CAP borrowers completed the baseline survey.

Renters sample. A total of 15,934 potential renters sample members were selected via random digit dialing from the 30 metropolitan statistical areas (MSAs) with the greatest representation of owners sample members. The dialing zones were initially restricted to the census block groups in which owners sample members were located but were each subsequently expanded to encompass a four-mile radius for those cases in which a more localized match with an owners sample member could not be obtained. Potential renters sample members were screened for tenure status and a household income ceiling. The income ceiling for each potential renters sample member was based on the area median income (AMI) of the matched owners sample member's MSA and the percentage minority representation of the matched owners sample member's census tract. The income threshold was set equal to 80% of the AMI if minority representation was less than 30%, or equal to 115% of the AMI if minority representation was 30% or greater. Calling and screening continued until the target baseline sample size had been achieved. The baseline renters sample comprises 1,529 cases.

Demographic Characteristics. The owners sample members were, on average, 35 years old with an annual household income of about $32,000 at the time of the baseline survey. Slightly more than half of the owners sample respondents (54%) were male, and about 46% were married at baseline. About half of the owners sample respondents also reported children in the household at baseline. Whites make up about 62% of the sample, followed by Blacks (19%) and Hispanics (16%). Approximately 51% of owners sample members had completed a high-school education as of baseline, while an additional 14%, 18%, and 6% had completed an associate's degree, bachelor's degree, or graduate degree, respectively. The labor force participation rate of the owners sample was about 96%, and approximately 92% of owners sample members were employed at baseline. In comparison, the renters sample members at baseline were slightly older (40) on average, had a somewhat lower average income ($20,200), and were more likely to be female (70%). Blacks (33%) and Hispanics (19%) have greater representation in the renters sample, and Whites (44%) have lower representation. The renters sample members were less likely to be married (27%) at baseline and were less likely to report children in the household (43%). The renters sample members also exhibited a lower rate of postsecondary degree completion (25%), labor force participation (75%), and employment (63%) at baseline.

Geographic Coverage. A majority of owners (62%) and renters (74%) sample members were located in the South at baseline. An additional 26% of the owners sample and 14% of the renters sample members were located in the Midwest, while 10% of the owners sample members and 12% of the renters sample members were located in the West. Only about 3% of the owners sample members and none of the renters sample members were located in the Northeast. At the state level, North Carolina accounts for 27% of the owners sample and 33% of the renters sample. Ohio accounts for an additional 12% of the owners sample and 6% of the renters sample, while Oklahoma contributes an additional 11% of the owners sample and 22% of the renters sample. Each of the other states represents less than 10% of both samples.