Internet usage and rural entrepreneurship in China

Chuanchuan Zhang
School of Economics
Central University of Finance and Economics
Email: ccz.zhang@gmail.com

Jianmei Zhao
China Academy of Public Finance and Public Policy
Central University of Finance and Economics
Email: jianmei@cufe.edu.cn

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

Internet usage and rural entrepreneurship in China

Abstract: This paper examines the impact of internet usage on rural entrepreneurship and investigates its possible mechanisms. We apply both two stage least square and bivariate probit regressions to identify the causal effects of internet usage on rural entrepreneurial activities. Based on the data from China Family Panel Studies, we find that internet usage significantly promotes rural entrepreneurial activities in China. Further evidence from subgroup regressions points towards greater impact of entrepreneurial behaviors for male relative to that for female. For younger and better educated people, internet usage imposes significant positive effects on their entrepreneurship, whereas such effects on older and less educated people is not remarkable. Internet usage eases the rural entrepreneurship via the channels of weakening financial constraint, lessening the role of social interaction and diminishing the restriction of transaction cost.

Keywords: Internet usage, rural entrepreneurship, channels, China

JEL: L86; O12; O33; R23

1. Introduction

Entrepreneurship or self-employment is important for economic growth, particular for transition economies, such as China, where majority of people live in rural areas, and Chinese government faces great challenges in providing jobs for vast rural labors. Rural entrepreneurial enterprises represent the most dynamic force in the Chinese economy (Pooh, Zhou, and Chan, 2009), and rural entrepreneurs are of extreme importance in China's progress toward a more market-oriented economy (Yu et al.,

2013). Nonetheless, starting and growing a business anywhere is fraught with multiple difficulties, particularly in rural areas that characterized by the imperfect credit market, supply chains and product markets (Banerjee and Newman, 1993).

With the rapid development of information technology, internet has gradually penetrated into rural societies. Reports from China internet network information center (CNNIC, 2017) indicate that, until the end of 2016, more than 200 million rural Chinese have accessed the internet, accounting for about 1/3 of rural population in China. The advent of internet undoubtedly promotes entrepreneurial conditions by facilitating entrepreneurial finance (Liu, 2015), building and maintaining social interaction (Pénard and Poussing, 2010), and reducing the transaction cost (Bauernschuster et al, 2014). On the other hand, internet may also discourage entrepreneurship by improving job search and increasing earnings (Pope, and Kroft, 2012) and providing the highest expected net income (Evans and Jovanovic, 1989). Which of the two opposing forces dominates? The theoretical ambiguity of the impact of internet on self-employment and the importance of entrepreneurship for sustained economic development in rural China inspire us to examine the effects of internet usage on rural individual's entrepreneurial activities.

The primary objective of this paper is to analyze the impact of a new information technology, internet, on individual's occupation choices. In particular, we investigate the causal effects of internet usage on rural individual's entrepreneurial activities using Chinese data. As a second objective, we investigate the possible channels through which internet usage affects rural entrepreneurship. Since more and more rural residents will use internet in less developing areas over the next few decades, this analysis may provide insights for developing countries to understand the changing nature of the labour markets with the coming of internet era.

Evaluating the effects of internet usage on entrepreneurship is not straightforward. Endogeneity problem originated from omitted variables and reverse causality between internet usage and self-employment may bias the estimate of internet variable. In this paper, we adopt both two stage least square (TSLS) regression and bivariate probit (Biporbit) regressions to identify the impact of internet usage on rural individual's entrepreneurship. We consider whether a rural household could access to natural gas pipeline for daily cooking as the IV in the baseline regression. The rationale for such an IV choice is that a rural household's access to the natural gas pipeline reflects its infrastructure level of energy supply that correlates with the household's other critical infrastructures, such as transportation, road infrastructure, water supply, and telecommunications. A households' access to the natural gas pipeline corresponds to its better telecommunication infrastructure that increases the probability of internet usage. However, the household's access to the natural gas pipeline is dependent on the decisions of energy companies, it is less likely to be correlated with an individual's entrepreneurial decision.

Based on above IV, both TSLS and Biporbit regressions present a significant positive effect of internet usage on rural entrepreneurship. The heterogeneous analysis indicates that, the impact of internet usage on individual's entrepreneurial behavior is more pronounced for male relative to that of female. For younger and better educated people, their use of internet creates significant positive effects on their entrepreneurial activities, whereas such effects on older and less educated people is not remarkable.

To ensure the causal impact of internet usage on rural entrepreneurial behavior, our first robustness check is to add province dummies to control for the effects of cultural preferences or other locational differences on entrepreneurship activities. Since finding an appropriate instrument in cross-sectional data is always a challenge, our second

robustness check is to construct another IV from the "supply side" of internet usage. We aggregate the number of telecom service outlets at the county level to represent local telecommunication infrastructure. We believe that greater number of telecom service outlets implies a better telecommunication infrastructure and greater possibility of internet usage for local residents. But the establishment of telecom service outlets is based on separate decisions of different telecom companies, and it is less likely to be correlated with an individual's entrepreneurial activities. To do this, we collect and match the aggregated number of telecom service outlets at the county level into the China Family Panel Studies (CFPS) data. Results from above robustness checks further confirm the positive effects of internet usage on rural entrepreneurship.

Having established that internet usage increases self-employment in rural China. We turn to investigate the possible mechanisms through which these effects operate. Results show that internet usage eases the rural entrepreneurship via the channels of weakening financial constraint, lessening the role of social interaction and diminishing the restriction of transaction cost.

The remainder of the paper is organized as follows. Section 2 summarizes the data and illustrates model specifications. In section 3, we present baseline regression results and heterogeneous effects of internet usage on rural entrepreneurship. We also conduct robustness checks in this part. In section 4, we investigate the possible channels through which that internet usage affects rural entrepreneurial activities. Section 5 concludes.

2. Model and data

2.1 Model and identification

Two possible endogeneity issues may complicate our estimation. The first complication arises because omitted variables may impact both internet usage and an

individual's entrepreneurial activity. For example, unmeasurable personal characteristics, such as less risk aversion or strong preferences in experiencing new technology innovation, could contemporaneously impact both entrepreneurial dicisions and internet usage of an individual. Ignorance of such unobservable determinants may invalidate our causal interpretation of internet usage on rural entrepreneurship.

The second complication arises because internet usage and entrepreneurial activities may suffer from a reverse causality. Internet exploration enables an individual to access to more information and find potential investment opportunities. Therefore, internet usage may encourage an individual to become an entrepreneur. Meanwhile, it is also plausible that the self-employed business owners have more incentive to seek business opportunities and understand market conditions. Internet, as an efficient information technology, is more likely preferred by entrepreneurs. As a consequence, it is the entrepreneurial activity that leads to the use of the internet. In either way, the ordinary least square regression will bias the estimate of internet usage.

The general solution to above endogeneity problem is the instrumental variable (IV) approach by running a two stage least square estimation. Considering that the dependent variable and the interested variable are both binary, in our case, we estimate a bivariate probit model (Biprobit) in which equations for the probability of entrepreneurship and the probability of the internet usage, are estimated simultaneously. The Biprobit model is equivalent to an instrumental variable or two-stage least squares model and it is preferred when both the dependent variable and endogenous variable are binary (Greene, 2003).

Assume that a self-employed business owner is identified by an unobserved latent variable, $Entrepreneur_i^*$ and that $Internet_i^*$ is the latent variable measuring an

individual's use of internet. Because these variables are not directly observable, we specify the model as follows:

$$Internet_i^* = \beta Z_i + \varepsilon_i$$
$$Internet_i = \begin{cases} 1, & \text{if} \quad Internet_i^* > 0 \\ 0, & otherwise \end{cases}$$

where Z_i refers to the observed determinants of internet usage, β is the associated parameters, and ε_i is a random error term. Analogously, the propensity to become an entrepreneur is measured by:

$$Entrepreneur_i^* = \delta\ Internet_i + \gamma X_i + u_i$$
$$Entrepreneur_i = \begin{cases} 1, & \text{if} \quad Entrepreneur_i^* > 0 \\ 0, & otherwise \end{cases}$$

where $Internet$ is our interested variable, and δ measures how internet usage affects rural entrepreneurial activities. X_i refers to the observed exogenous determinants of being an entrepreneur, γ is the associated parameters, and u_i is a random error term. Yet because the two variables are potentially explained by common determinants, the error terms of the two models are dependent and distributed as a bivariate normal, so that $E(\varepsilon_i) = E(u_i) = 0$, $\text{var}(\varepsilon_i) = \text{var}(u_i)=1$, and $\rho = \text{cov}(\varepsilon_i, u_i)$. A Wald test for $\rho = 0$ indicates whether the models should be jointly estimated.

In practice, the choice of Z_i is difficult. We need to find an IV that affects the use of internet, meanwhile it should be uncorrelated with an individual's entrepreneurship activities. In this application, one potential exclusion restriction is the situation on household domesticity infrastructure. In particular, we create a dummy variable that equals one if the rural household could access to the natural gas for its daily cooking. Like most other countries that there exists a growing trend of mutual reliance among the economic critical infrastructures (Schneider et al, 2006; Svendsen and Wolthusen,

2007; Ouyang, 2014), economic infrastructures [1] in rural China, including transportation, road infrastructure, piped water supply, piped gas supply, and telecommunication are also mutually interdependent (Wang et al., 2012). Rural households' access to the natural gas pipeline implies a better infrastructure of energy supply, which should correspond to a better infrastructure of telecommunication that enables the households to access and use the internet. The households' access to the natural gas pipeline should satisfy two necessary properties of a valid instrumental variable —it reflects the critical infrastructure level, including telecommunication level that affects the rural household's likelihood of internet usage. However, we do not expect the household's connection to the gas pipeline has a strong direct effect on an individual's entrepreneurial decision.

2.2 Data

Our empirical analysis is mainly based on the 2014 China Family Panel Studies (CFPS), which was conducted by the Institute of Social Sciences Survey Center at Peking University, China, in collaboration with the Survey Research Center at University of Michigan. The survey was designed in a style similar to the Panel Study of Income Dynamics (PSID) in the U.S. The CFPS covers a wide range of domains for families and individuals from 162 counties in 25 provinces of China; it is a nationally representative household survey that contains rich socioeconomic information for individuals, families and communities (Xie and Hu 2014). The latest 2014 CFPS survey collects information on individual's internet usage behavior, which provides a unique feature to examine the effects of internet usage on entrepreneurship in China.

[1] Economic infrastructures include public utilities (i.e., power, telecommunications, piped water supply, sanitation and sewerage, solid waste collection and disposal, and piped gas), public works (i.e., roads and major dam and canal works for irrigation and drainage), and other transport sectors (i.e., urban and interurban railways, urban transport, ports and waterways, and airports). See the work bank (1994). A broad definition of infrastructure also includes human capital (e.g., education and health).

To assess the potential mechanisms that internet usage affects rural entrepreneurship, we additionally collect data at the county-level concerning financial institutions and telecom service outlets to measure the financial availability and telecommunication infrastructure, and merge above county-level variables into the CFPS covered areas. The total number of financial institutions within each county/district is from China Banking Regulatory Commission. The aggregated number of telecom service outlets is obtained from the largest three telecom companies in China, China Telecom, China Mobile and China Unicom. The three major telecom companies dominate the whole telecom market, accounting for about 99% of telecommunication market and telecom business. Telecom service data were obtained from headquarters of the three largest telecom operators, respectively.

Table 1 presents definition, means and standard deviations of key variables in the analysis. The full sample size is 11604. We classify a person as an entrepreneur if his/her employment status is either "self-employment" or "employer", and his/her working place is "private enterprise" or "individual industrial and commercial household". According to this definition, rural entrepreneurs account for 8.4% of the sample population. Internet usage is a binary variable. In our sample, about 26% rural adults are internet users. This number is reasonable and matches to the national report from CNNIC, that the percentage of rural internet users (including teenagers) was 27.5% at the end of year 2014. Variables at the individual level include age, gender, marital status, self-rated health status and personal education category.

At the household level, we include family size, household wealth measured with a dummy variable indicating whether the household owns a second tradable commercial housing, besides its countryside homestead. It is worth noting that, despite housing is the dominant assets for both urban and rural families, the housing systems for urban

and rural families are quite different after the housing reform in China. Unlike urban housing that can be traded on the housing market, rural housing built on the countryside homestead (distributed from the Rural Collective) suffers stringent restrictions for commercial transaction, rural real estate and homestead are rarely traded on the property market. As a consequence, wealthier rural households purchase commercial real estates to accumulate their assets. We also include land variable measured with binary variable to show whether the rural household owns rural collective land. The instrumental variable indicating whether the household could access to the natural gas for its daily cooking is also binary. This variable is used to reflect the household living infrastructure. In addition, we use money amount of gift expenditure to measure the social interaction of the household. We also include the distance from household living community to the nearest town to approximate the transaction costs for local rural businesses.

The county-level variables include the total number of financial institutions that measures financial availability and the aggregated number of telecom service outlets that reflects telecom infrastructure. Those data are obtained from outside data source, they are merged into the CFPS data.

[Table 1 here]

3. Results, heterogeneous effects and robustness

3.1 Baseline results

Considering the access to the natural gas pipeline as an instrumental variable, both two stage least square (TSLS) estimates and bivariate probit (Biprobit) estimates are presented in Table 2. Table 2 provides estimated coefficients and robust standard errors for the variables clustered at the household level.

[Table 2 here]

The coefficient of the instrumental variable, the access to the natural gas pipeline that represents a better local infrastructure, shows a significant positive sign in the first stage estimation of TSLS. The internet variable also imposes significant effects on individual's entrepreneurial activities, therefore the internet usage positively impacts individual's occupation choice of being an entrepreneur. Similarly, the Biprobit regression shows that the household's access to the natural gas pipeline positively affects the probability of an individual's internet usage, and rural internet users are more likely to be a self-employed business owners. Overall, we could observe a positive effect of internet usage on rural individuals' entrepreneurial activities.

For controlled variables, both regressions show that the age variables have a concave relationship with the rural entrepreneurship. Marries individuals tend to be self-employed. Individuals within a larger family tend to carry out some entrepreneurial activities. Rural households that are capable to afford a second residential property, their family members are more likely to be self-employed entrepreneurs.

3.2 Heterogeneous effects

To see whether the internet usage imposes the homogeneous effects on a rural individual's entrepreneurship, we conduct several subgroup analyses and report the results in Table 3.

[Table 3 here]

First, we compare the effects of internet usage on the occupation choices between genders. Results in column 1 and 2 show that for male, internet usage has much larger and stronger positive impact on entrepreneurial activities of male than that of female. This is consistent with Rijkers and Costa (2012), who found that women are less likely to be non-farm entrepreneurs than men are.

Since internet is a newly developed telecom technology that requires literacy and information processing ability, we conduct subgroup analyses by age and education levels separately. Results in column 3-4 show that the effects of the internet usage on entrepreneurial activities is significant for young adults, whereas its effects on adults at older age is not remarkable. The last 5-6 columns in Table 3 show that internet usage significantly encourages the entrepreneurial activities for rural individuals with higher education level, whereas the effects of internet usage on individuals with lower education are not significant. On the whole, although the internet usage plays a positive role in rural individuals' self-employed occupation choices, its impact on different groups is heterogeneous.

3.3 Robustness checks

As shown in the Table 2, both TSLS and Biprobit regressions reveal that internet usage promotes an individual's entrepreneurial activities in rural China. One concern might be that there exists certain time-constant cultural preferences or other locational differences for entrepreneurship activities. Therefore, we add province dummies in the regression to control such effects approximately. The results are reported in the column 1 - 2 of Table 4. The effects of internet usage on the probability of being an entrepreneur remains after controlling province dummies in the regression.

[Table 4 here]

Another concern is the challenge that considering the access to the natural gas pipeline as the IV to proximate the level of telecom infrastructure. Despite we believe the validity of such IV, we employ another exclusion restriction at the county level, the total number of telecom service outlets from China's largest telecom companies (China Telecom, China Mobile, and China Unicom). The aggregate number of telecom service outlets within each county reflects the level of local telecommunications infrastructure.

More telecom service outlets imply a better telecommunication service and greater probability of internet usage for local residents. However, the aggregated number of telecom service outlets is dependent on the three telecommunication companies separately, it is less likely to correlate with an individual's entrepreneurial activities.

The Biprobit regression results based on the new instrumental variable, the number of telecom service outlets, are shown in column 3 - 4 of Table 4. Again, by employing the number of county-level telecom service outlets as the IV and adding the province controls in the regression, internet usage continues to remain the positive effects on rural individual's entrepreneurial behaviors.

4. Possible channels

Having established that internet usage eases a rural individual's entrepreneurial activities. We turn to investigate the possible channels through which internet usage promotes the occupation choices of becoming an entrepreneur.

4.1 Internet usage and financial availability

There is a large literature suggests that finance is one of the most important factors for entrepreneurship. Financial constraint creates barriers to most entrepreneurs. Imperfections in the credit market could deter potential entrepreneurs to acquire necessary capital to start a business (Evans and Jovanovic, 1989; Paulson and Townsend, 2004; Bianchi and Bobba, 2013). Formal bank credit availability not only raises the rate of business incorporations (Black and Strahan, 2002), but also promotes personal entrepreneurship (Karaivanov, 2012).

The evolution and recent developments of information and communication technology has changed the way organizations operate and do business especially in the banking industry. The introduction of E-banking has impacted on the efficiency and

effectiveness of banking service, resulting in the improvement of the bank-customers relationship as well as customer satisfaction (Ojokuku and Sajuyigbe, 2015). Furthermore, the latest developed financing alternatives, such as crowdfunding and peer-to-peer lending originated from the internet, provide novel financing options for entrepreneurs to start and grow ventures (Bruton et al., 2014). Relative to the non-users, it is much easier for internet users to access to the financial resources and unleash the financial constraint, particular for self-employed business owners.

To empirically gauge the impact of financial availability on internet users versus non-users, we employ a measure of total number of total bank branches at the county level as the proxy for financial availability for potential entrepreneurs. We also interact the internet usage variable with the number of bank branches in the second equation of bivariate probit regression.

$$\Pr(Entrepreneur_i = 1 \mid X) = \lambda_1 Internet_i + \lambda_2 FinAvail_i + \lambda_3 (Internet_i \times FinAvail_i) + X_i \gamma_1$$

The regression outputs are presented as channel I in Table 5.

[Table 5 here]

By including all other control variables in the previous regression, it shows both internet usage and financial availability positively affect people's entrepreneurial activities ($\lambda_1 > 0, \lambda_2 > 0$). The interaction term between internet usage and financial availability measure is negative and statistically significant ($\lambda_3 < 0$), suggesting that, for internet users, the effects of financial threshold has been reduced relative to that of non-users.

4.2 Internet usage and social interaction

The second possible channel that we investigate is the social interaction that relates to the social capital. Social capital has been argued to matter for a variety of economic outcomes, in particular in the presence of asymmetric information and incomplete contracts. To name but a few examples that relate directly to individual networks, social

capital is found to be relevant for job search (Mouw, 2003; Bayer et al., 2008), entrepreneurial activities (Sanders and Nee, 1996; McMillan and Woodruff, 1999; Westlund and Bolton, 2003; Bauernschuster, et al., 2010; Kwon et.al., 2013), firm location (Michelacci and Silva, 2007), and firm innovation (Molina Morales and Martínez Fernández, 2010).

Traditionally, rural enterprises rely heavily on social capital in rural communities. The advent of internet has changed people's communication style, resulting in a reduction of people's face-to-face communication. Nonetheless, recent research finds that, instead of undermining social capital, internet usage is actually a new way of communication and socialization that can supplement inter-personal relations and civic engagement. Internet has become a convenient and efficient means of maintaining existing social ties and/or of creating new ties (Pénard and Poussing, 2010; Bauernschuster, et al., 2014). Even in rural communities, internet usage could help in facilitating social and business communication.

To investigate how internet usage affects rural entrepreneurs through the social interaction, we measure personal social interaction with the annual gift expenditure at the household level. Such social capital "investment" reflects one's social networking, or "Guanxi" in rural communities. Such social interaction enables enterprising individuals to draw on these resources when establishing and operating small businesses. In the similar vein, we interact internet usage with social capital variables in the Biprobit regression and present results as channel II in Table 6.

$$\Pr(Entrepreneur_i = 1 \mid X) = \beta_1 Internet_i + \beta_2 SocialCap_i + \beta_3 (SocialCap_i \times Internet_i) + X_i \gamma_2$$

[Table 6 here]

As expected, both internet usage and social interaction variables show positive effects on rural entrepreneurial activities ($\beta_1 > 0, \beta_2 > 0$). Meanwhile, the coefficient on the

interaction term appears as significant negative ($\beta_3 < 0$). The combined estimation results suggest that promoting effects from internet usage on rural entrepreneurship weakens in rural communities with stronger social interaction, relative to that from communities with less social interaction.

4.3 Internet usage and transaction cost

More plausibly, internet usage may have lowered the transaction cost, thus stimulating rural self-employment and development of micro enterprises (Lohrke et al., 2006; Wu et al., 2014). Typical rural markets are characterized by asymmetric information, in which market information searching can be very costly. First, market information searching for products sale or price of inputs entails substantial transportation and labour costs. Internet usage is an effective remedy for such information problem, saving transportation costs and raising transaction execution efficiency for owners of small business. Second, rural entrepreneurs could also benefit from internet usage by increasing their geographic reach or acquire new customers to expand their business scope, the scale economy may also reduce rural SME's transaction costs.

Empirical measurement of the transaction cost for rural small businesses is unavailable, due to the reason that CFPS are unable to provide direct transaction costs data from self-employed business owners. We use the distance from the household to the nearest business market to proxy the transaction cost of rural entrepreneurs. We introduce the interaction term of internet and transaction cost variables in the Biprobit regression, and regard it as another possible channel that internet usage affects rural entrepreneurship. Regression results are outputted in Table 7.

$$\Pr(Entrepreneur_i = 1 \mid X) = \alpha_1 Internet_i + \alpha_2 TranCost_i + \alpha_3 (TranCost_i \times Internet_i) + X_i \gamma_3$$

[Table 7]

Internet usage still remains a positive effect on entrepreneurial activities ($\alpha_1 > 0$), whereas transaction cost variable imposes negative effects on rural entrepreneurship ($\alpha_2 < 0$), implying that higher transaction cost may discourage rural entrepreneurship. Finally, the interaction term of internet usage and transaction cost shows significant positive sign ($\alpha_3 > 0$).This means that internet usage has greater effects in encouraging entrepreneurship for rural area entails greater transaction cost than that with less transaction cost.

5. Conclusion

This article presents some evidence on the impact of a new information technology, internet, on rural labor occupation choices in China. We not only examine the effects of internet usage on rural entrepreneurship, but also investigate the possible mechanisms through which these effects operate.

Based on the CFPS data in China, we address the endogenous problem with both two stage least square and bivariate probit regressions. Results show that the internet usage of rural residents significantly eases their entrepreneurial activities. Further evidence from subgroup regressions points towards greater impact of entrepreneurial behaviors for male relative to that for female. Additionally, we find that the internet usage imposes significant positive effects on the entrepreneurship of younger and better educated people, whereas such effects on older and less educated people is not remarkable. After established that internet usage promotes a rural individual's entrepreneurial activities. We next analyze the possible mechanisms and reveal that internet usage eases the rural entrepreneurship via the channels of weakening financial

constraint, lessening the role of social interaction and diminishing the restriction of transaction costs.

As more and more rural residents will become internet users in less developing areas over the next few decades, this analysis may provide insights for developing countries to understand the changing nature of the labour markets with the advent of internet era.

References

1. Bauernschuster, S., Falck, O., and Woessmann, L., 2014. Surfing alone? The Internet and social capital: Evidence from an unforeseeable technological mistake. Journal of Public Economics 117, 73-89.
2. Bayer, P., Ross, S. L., and Topa, G., 2008. Place of work and place of residence: Informal hiring networks and labor market outcomes. Journal of Political Economy 116, 1150-1196.
3. Bianchi, M., and Bobba, M., 2013. Liquidity, risk, and occupational choices. The Review of Economic Studies, s31.
4. Bourdieu, P., 1986. Handbook of Theory and Research for the Sociology of Education. Greenwood Press.
5. Xie, Y., and Hu, J., 2014. An introduction to the China family panel studies (CFPS). Chinese Sociological Review 47, 3-29.
6. Bruton, G., Khavul, S., Siegel, D., and Wright, M., 2015. New Financial Alternatives in Seeding Entrepreneurship: Microfinance, Crowdfunding, and Peer‐to‐Peer Innovations. Entrepreneurship Theory and Practice 39, 9-26.
7. Evans, D. S., and Jovanovic, B., 1989. An estimated model of entrepreneurial choice under liquidity constraints. Journal of political economy 97, 808-827.

8. Guiso, L., Sapienza, P., and Zingales, L., 2004. The role of social capital in financial development. The American Economic Review 94, 526-556.
9. Hurst and Lusardi, 2004; Paulson and Townsend, 2004; Djankov et al., 2006; Fairlie and Krashinsky, 2012; Wang, 2012).
10. Lohrke, F. T., Franklin, G. M., and Frownfelter-Lohrke, C., 2006. The Internet as an Information Conduit A Transaction Cost Analysis Model of US SME Internet Use. International Small Business Journal 24, 159-178.
11. McMillan, J., and Woodruff, C., 1999. Interfirm relationships and informal credit in Vietnam. The Quarterly Journal of Economics 114, 1285-1320.
12. Michelacci, C., and Silva, O., 2007. Why so many local entrepreneurs? The Review of Economics and Statistics 89, 615-633.
13. Molina Morales, F. X., and Martínez Fernández, M. T., 2010. Social networks: effects of social capital on firm innovation. Journal of Small Business Management 48, 258-279.
14. Mouw, T., 2003. Social capital and finding a job: do contacts matter? American sociological review, 868-898.
15. Ojokuku, R. M., and Sajuyigbe, A. S., 2012. The Impact of Electronic Banking on Human Resources Performance in the Nigerian Banking Industry. International Journal of Economic Development Research and Investment 3, 61-69.
16. Ouyang, M., 2014. Review on modeling and simulation of interdependent critical infrastructure systems. Reliability engineering & System safety 121, 43-60.
17. Paulson, A. L., and Townsend, R., 2004. Entrepreneurship and financial constraints in Thailand. Journal of Corporate Finance 10, 229-262.
18. Pénard, T., and Poussing, N., 2010. Internet use and social capital: The strength of virtual ties. Journal of Economic Issues 44, 569-595.

19. Rijkers, B., and Costa, R., 2012. Gender and rural non-farm entrepreneurship. World Development 40, 2411-2426.
20. Sanders, J. M., and Nee, V., 1996. Immigrant self-employment: The family as social capital and the value of human capital. American sociological review, 231-249.
21. Schneider, K., Liu, C., and Paul, J., 2006. Assessment of interactions between power and telecommunications infrastructures. IEEE Transactions on Power Systems 21, 1123-1130.
22. Sobel, J., 2002. Can We Trust Social Capital, Journal of economic literature, pp. 139-154.
23. Svendsen, N. K., and Wolthusen, S. D., 2007. Connectivity models of interdependency in mixed-type critical infrastructure networks. Information Security Technical Report 12, 44-55.
24. Wang, S., Hong, L., and Chen, X., 2012. Vulnerability analysis of interdependent infrastructure systems: A methodological framework. Physica A: Statistical Mechanics and its applications 391, 3323-3335.
25. Wu, L., Chen, K., Chen, P., and Cheng, S., 2014. Perceived value, transaction cost, and repurchase-intention in online shopping: A relational exchange perspective. Journal of Business Research 67, 2768-2776.

Table 1. Variable definition and summary statistics

Variables	Definitions	Mean	St. Dev.
Dependent variable			
Entrepreneur	Entrepreneur=1, if the employment status is either "self-employment" or "employer", 0 otherwise	0.084	0.277
Variables at the individual level			
Internet	=1 if the individual uses internet, 0 otherwise	0.258	0.438
Age	Adult age from 16 to 60 years old	39.927	12.721
Gender	=1 for male, ＝0 for female	0.488	0.500
Married	=1 for married status, 0 otherwise	0.808	0.394
Education	Illiterate=1 (28.039%); Primary=2(26.402%); Secondary=3(28.202%); High school and above=4(9.441%)	2.207	0.989
SRH	Self-rated health status, Excellent=1(17.9%), Very good=2(24.076%); good=3 (31.57%); fair=4(12.542%); poor=5(13.843%)	2.803	1.264
Variables at the household level			
Family size	Number of persons within the household	4.736	1.931
Wealth	Wealth=1 if the household owns at least two houses, 0 otherwise	0.139	0.346
Land	Land=1, implying that the rural household owns rural collective land, 0 otherwise	0.959	0.199
Access to gas pipeline	Gas=1, the household could access natural gas pipeline for daily cooking, 0 otherwise	0.221	0.415
SocialNetwork	Expenditure on gifts (or cash transfer) for remaining the social relations in past 12 months (yuan)	7.607	1.176
TranCosts	Logged distance from living community to the nearest town, approximating the transaction costs for rural business (0.5km)	9.326	12.462
Variables at the county level			
FinancialIns	Logged number of all financial institutions within the county/district, representing the degree of financial availability	6.388	0.535
Telesites	Logged number of telecommunication service outlets within the county/district, representing the telecommunications infrastructure	3.465	0.683

Sources: China Family Panel Studies 2014. The full sample size is 11604.

Table 2. The effects of internet usage on rural entrepreneurship

	Two stage least squares		Biprobit regression	
Dependent variable	1st stage Internet	2nd stage Entreneur	Pr(Internet=1|Z)	Pr(Entreneur=1|X)
GasPipeline	0.079***	-	0.801***	-
	(0.008)		(0.079)	
Internet		0.713***	-	1.717***
		(0.098)		(0.296)
Gender	0.025***	0.004	0.437***	0.291***
	(0.007)	(0.008)	(0.073)	(0.085)
Age	-0.045***	0.043***	-0.031	0.202***
	(0.002)	(0.006)	(0.028)	(0.034)
Age2	0.000***	-0.000***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Married	-0.055***	0.071***	-0.457***	0.447**
	(0.012)	(0.014)	(0.112)	(0.150
Primary	0.046***	-0.003	1.257***	0.688***
	(0.009)	(0.011)	(0.128)	(0.132
Secondary	0.152***	-0.038**	1.914***	0.823***
	(0.009)	(0.017)	(0.124)	(0.136)
High school and above	0.271***	-0.112***	2.742***	0.691***
	(0.012)	(0.028)	(0.146)	(0.179)
Self-ated health	0.006**	-0.001	0.062*	0.017
	(0.003)	(0.003)	(0.032)	(0.035)
FamSize	-0.005***	0.008***	-0.040 **	0.066**
	(0.002)	(0.002)	(0.019)	(0.021)
Wealth	0.043***	0.024**	0.305***	0.578***
	(0.009)	(0.011)	(0.095)	(0.095)
Land	-0.015	-0.021	-0.206	-0.556**
	(0.016)	(0.019)	(0.171)	(0.172)
Constant	1.356***	-1.164***	0.515	-7.934***
	(0.037)	(0.162)	(0.481)	(0.735)
Observations		9064		9064

Notes: "***", "**", "*" refer to statistically significant at 1%, 5%, and 10% level, respectively.

Table 3. The heterogeneous effects of internet usage on entrepreneurship, biprobit regressions on different subgroups

	Gender		Age		Education	
	Male	Female	Young	Old	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Internet	1.804***	1.207*	1.815***	1.243	2.103***	0.478
	(0.400)	(0.505)	(0.459)	(0.860)	(0.391)	(0.570)
Gender	-	-	0.235**	0.406**	0.108	0.677***
			(0.117)	(0.139)	(0.114)	(0.134)
Age	0.199***	0.182**	-	-	0.205***	0.153**
	(0.041)	(0.062)			(0.044)	(0.057)
Age2	-0.002***	-0.002***	-	-	-0.002***	-0.002***
	(0.000)	(0.001)			(0.001)	(0.001)
Married	0.554**	0.278	1.125***	-0.027	0.470**	0.470*
	(0.187)	(0.273)	(0.184)	(0.286)	(0.189)	(0.259)
Primary	0.195	1.055***	0.339	0.730***	-	-
	(0.178)	(0.209)	(0.232)	(0.184)		
Secondary	0.276	1.453***	0.469	0.808***	-	-
	(0.179)	(0.220)	(0.266)	(0.191)		
High school and above	0.284	0.956**	0.181*	0.702**	-	-
	(0.222)	(0.320)	(0.351)	(0.255)		
Self-rated health	0.004	0.030	0.101*	-0.062	0.057	-0.059
	(0.046)	(0.058)	(0.050)	(0.051)	(0.050)	(0.051)
FamSize	0.032	0.035	0.050*	-0.036	0.062*	-0.018
	(0.027)	(0.037)	(0.028)	(0.036)	(0.029)	(0.034)
Wealth	0.505***	0.495***	0.404	0.669***	0.304*	0.898***
	(0.131)	(0.165)	(0.141)	(0.156)	(0.131)	(0.157)
Land	-0.385	-1.116***	-0.559*	-0.801**	-0.319	-1.210***
	(0.262)	(0.274)	(0.244)	(0.296)	(0.265)	(0.257)
Constant	-7.671***	-7.725***	-4.747***	-17.324	-8.003***	-21.060
	(1.193)	(1.851)	(0.744)	(1277.81)	(1.168)	(645.500)
Province dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4653	4411	3658	5136	3661	5403

Notes: "***", "**", "*" refer to statistically significant at 1%, 5%, and 10% level, respectively.

Table 4. The effects of internet usage on rural entrepreneurship, with a different instrumental variable and province dummies controlled (Biprobit regressions)

	IV= Home gas cooking with province dummies		IV= Telecommunication sites with province dummies	
Dependent variable	Pr(Internet=1|Z)	Pr(Entreneur=1|X)	Pr(Internet=1|Z)	Pr(Entreneur=1|X)
GasPipeline	0.708***	-	-	-
	(0.091)			
Telesites	-	-	0.335***	-
			(0.072)	
Internet		1.377***		0.819**
		(0.321)		(0.384)
Gender	0.445***	0.321***	0.444***	0.364***
	(0.074)	(0.087)	(0.075)	(0.089)
Age	-0.056*	0.183***	-0.055*	0.155***
	(0.028)	(0.035)	(0.029)	(0.037)
Age2	-0.001**	-0.002***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Married	-0.471***	0.474**	-0.450***	0.473***
	(0.116)	(0.155)	(0.117)	(0.160)
Primary	1.179***	0.628***	1.143***	0.642***
	(0.134)	(0.136)	(0.135)	(0.141)
Secondary	1.798***	0.824***	1.750***	0.898***
	(0.130)	(0.140)	(0.132)	(0.147)
High school and above	2.611***	0.711***	2.571***	0.850***
	(0.152)	(0.182)	(0.153)	(0.195)
Self-rated health	0.047	0.008	0.045	0.016
	(0.033)	(0.036)	(0.033)	(0.037)
FamSize	-0.031	0.033	-0.044**	0.033
	(0.021)	(0.022)	(0.021)	(0.022)
Wealth	0.318**	0.525***	0.333***	0.567***
	(0.099)	(0.102)	(0.100)	(0.103)
Land	0.045	-0.691***	0.060	-0.744***
	(0.187)	(0.188)	(0.192)	(0.189)
Constant	1.547*	-7.653***	0.923	-6.644***
	(0.688)	(1.013)	(0.726)	(1.088)
Province dummies	Yes	Yes	Yes	Yes
Observations		9064		8916

Notes: "***", "**", "*" refer to statistically significant at 1%, 5%, and 10% level, respectively.

Table 5. Internet usage affects rural entrepreneurship via financial availability

Channel I	Estimates	St. Dev.
Channel I		
IntNet	4.859***	1.001
Financial availability	0.328***	0.096
IntNet * Financial availability	-0.476***	0.146
Controls	Yes	

Notes: "***", "**", "*" refer to statistically significant at 1%, 5%, and 10% level, respectively.

Table 6. Internet usage affects rural entrepreneurship via social interaction

Channel II	Estimates	St. Dev.
IntNet	2.853***	0.690
Social interaction	0.140***	0.049
IntNet * Social interaction	-0.142*	0.077
Controls	Yes	

Notes: "***", "**", "*" refer to statistically significant at 1%, 5%, and 10% level, respectively.

Table 7. Internet usage affects rural entrepreneurship via transaction cost

Channel III	Estimates	St. Dev.
IntNet	1.540***	0.321
Transaction cost	-0.017**	0.007
IntNet * Transaction cost	0.015*	0.009
Controls	Yes	

Notes: "***", "**", "*" refer to statistically significant at 1%, 5%, and 10% level, respectively.