Does Fintech Credit Reduce Income Inequality? Evidence from Migrant versus Native Business Owners

Pan Jiang *
Wilfrid Laurier University

Si Li
Wilfrid Laurier University

Oct 2, 2021

Abstract

We examine whether Fintech-based microfinance reduces the income gap between the migrant and the native small business owners. Using the data on business owners registered with the largest Fintech firm in China, we find that with the access to Fintech microloans, migrants achieve greater business revenues compared to their native counterparts. The effect is more pronounced in the businesses with more financial constraint, in more economically developed areas, for more risk averse business owners, and among the owners who are relatively new to the Fintech platform. Overall, our findings support that Fintech plays an important role in reducing income gap between native and migrant microentrepreneurs.

Keywords: Fintech, digital finance, microfinance, migrant, income inequality

---

* Pan Jiang (pjiang@wlu.ca), Financial Services Research Centre, School of Business and Economics, Wilfrid Laurier University, Waterloo, Ontario N2L 3C5, Canada. Si Li (sli@wlu.ca), Financial Services Research Centre, School of Business and Economics, Wilfrid Laurier University, Waterloo, Ontario N2L 3C5, Canada. The authors thank Ant Group and its "Peak Initiative of Digital Finance Open Research" for generously providing the data used in this paper. Both Li and Jiang acknowledge financial support from the Lazaridis Institute Grant.

1. Introduction

Income inequality is an important topic to academics, politicians, and practitioners. The widening of income inequality has remained a major political issue and a great concern for economic growth (Cingano, 2014). The growth in financial technologies (Fintech) in recent years could offer one candidate solution to reduce income inequality. Both the United Nations 2030 Agenda for Sustainable Development and the G20 High-Level Principles for Digital Financial Inclusion highlight the importance of harnessing Fintech to reduce financial exclusion and income inequality.[1] By reducing financial market imperfections such as information asymmetries and transaction costs, Fintech allows the disadvantaged group to access the financial market that is traditionally not accessible to them (Galor and Zeira, 1993). Empirically, the literature finds a significant impact of Fintech on reducing poverty (Jack and Suri, 2016) and income inequality (Demir, et al., 2020). Our paper adds to the growing body of the literature on the inclusive role of Fintech, and provides the evidence that digital loans and mobile payments enabled by Fintech help reduce income gap.

This paper investigates the income gap between migrant and native workers. Specifically, we argue that Fintech reduces the obstacle to migrants' access to the financial market and thus lessens the migrant-native income gap. Migrants are often disadvantaged compared to native citizens. For example, existing evidence shows that migrant workers suffer substantial discrimination in the labor market (Friedman and Lee, 2010) and have lower average incomes (Knigh et al., 2010, Meng and Bai, 2007). Using the data from

[1] UNSGSA: Igniting SDG Progress Through Digital Financial Inclusion:
https://sustainabledevelopment.un.org/index.php?page=view&type=400&nr=2655&menu=1515
G20 High-Level Principles for Digital Financial Inclusion:
https://www.gpfi.org/publications/g20-high-level-principles-digital-financial-inclusion

China, Li (2010) and Wang and Tian (2014) find that migrants' lower socioeconomic status together with the household registration restrictions creates barriers for migrants to access financial services in China, such as obtaining a mortgage. As a result, Fintech may help decrease the barrier for migrants to access the local financial services and markets, and increased accessibility may narrow the income gap between migrants and their native counterparts.

The income gap between migrants and natives has important economic implications. Migrations result in labor force reallocation and improve economic efficiency. Based on Ratha, et al. (2011), migration often happens for economic reasons. Labor tends to move from geographical areas or industry sectors with lower return of labor to those with higher return, improving the overall efficiency as well as reducing income inequality among people who are originally from different regions. Hao, et al., (2020), for example, use the data from China and show that migration-induced labor reallocation plays a central role in the country's aggregate income growth and regional income convergence.

The income gap between migrants and natives is also of interest to policy makers. Narrowed income gap not only indicates a higher degree of equality and integration for existing migrants but also provides a catalyst for potential future migration and urbanization. Specifically derived from our paper, Fintech enables a better access to financial markets, and thus provides more opportunities for migrants to run small businesses. These businesses serve as an alternative job opportunity for migrants who suffer from discrimination in the traditional job market. Given that migration and inequality faced by migrants in destination cities are a global phenomenon, the implication of our research, even if based on the China data, goes beyond the scope of China.

Despite the important economic and policy implications, the research on how financial market and Fintech impact migrants and income inequality remains largely unexplored. Datta (2009) investigates financial exclusion among migrants in London, England and finds migrants' migration history, immigration status and transnational lives affect their perceived risk to use financial service. Anderloni and Vandone (2006) analyze migrants' demand for financial services and calls for financial innovation to satisfy such demand. Albareto and Mistrulli (2011) find that migrants pay, on average, almost 70 basis points higher for credit than natives do. Recently, Fintech-enabled innovations that provide more inclusive services to unprivileged people (e.g., the unbanked low-income households) are receiving increasing attention. For example, Jack and Suri (2016) find a meaningful reduction of poverty resulting from mobile money usage. Demir, et al. (2020)'s study of 140 countries find that Fintech reduces overall income inequality measured by Gini coefficients. To summarize, the research on migrants largely focuses on the migration decision of the labor force and the well-being of the migrant households. We extend this literature by examining how Fintech innovations affect the income of migrants compared to that of their native counterparts in the same city.

The difficulty in investigating income inequality is the availability of individual-level income data because individual workers' wage data is often not available. Extant studies use survey data, e.g., Current Population Survey, to investigate the workers' earnings and other characteristics.[2] In this paper, we use the individual business owners' revenue and loan access data provided by Ant Group, a subsidiary of Alibaba that

[2] The Census Bureau may provide confidential and proprietary data on individual workers' earnings. However, the identification of migrants vs. natives cannot be easily done using the Census data. In addition, the access to the U.S. financial markets for migrants is not distinctly different for natives.

specializes in online financial services. Our sample consists of individuals registered with Ant Group using their personal information (i.e., personal identification card number) rather than corporate information. The medium monthly income of our sample is CNY5,253.00 (USD763.68), suggesting that a typical individual in our sample runs a micro business.[3] We examine the income gap between migrant and native business owners based on their business revenues from their micro businesses registered with Ant Group. Because many migrants do not have formal jobs and their wage income is not available from the China Census Bureau, their small business revenues serve as a reasonable proxy for their income. One caveat is that it is possible that our sample business owners may have income sources other than the business they run through the platform provided by Ant Group. The result in this paper thus needs to be interpreted with this caveat in mind.

We use the difference-in-difference (DID) approach to compare the impact of drawing a Fintech-based microloan on business revenues for migrants with that for natives. Our main finding is that migrant business owners with the access to microloans harvest more revenues compared to their native counterparts. This suggests that Fintech-based microfinance benefits migrants more than it does natives because Fintech-based microfinance reduces the obstacle resulting from financial exclusion to migrants while such an obstacle is likely not existing to native business owners. Our result thus provides the evidence of Fintech reducing the migrant-native income inequality.

Heterogeneity analysis shows that the incremental benefit migrants obtain from the Fintech loan exists in the businesses with more financial constraint, in more economically developed areas, for more risk averse business owners, and among the owners who are

[3] We convert Chinese Yuan (CNY) into U.S. dollars based on the exchange rate on December 31, 2018, 1USD = 6.8785CNY.

relatively new to the Fintech platform. We first employ triple DID to explore the heterogeneity among individuals with more versus less financial constraint. We employ two measures to identify more financially constrained individual business owners. First, financial constrained micro business owners are defined as those who have lower revenue (where "lower" is identified by below median, below first quartile or below first tercile) at the beginning of our sample, March 2017.[4] Second, we classify "more financially constrained" business owners as those having QR money-receiving code but no online shop. Operating an online shop typically requires more resources (requiring at least a laptop plus internet connection) than operating offline using a QR code (only requiring a smart phone with data subscription). We find that the beneficial effect to migrants is more prominent for the business owners with lower initial revenue and those who have no online shops.

We then divide our sample based on economic development levels: city GDP per capita (high, medium and low); and geographic location (south vs. north; east vs. non-east), where southern and eastern area are more economically advanced. We find the effect of migrants obtaining incremental benefit relative to natives after drawing a Fintech microloan is consistently present in more developed cities or regions. Next, we split the sample by individual business owners' gender (male vs. female) and age (younger and elder based on age median), and we find the effect only exists in female subgroup and elder subgroup, both are typically considered as more risk averse population. We finally divide the sample based on signup months for the Fintech platform, and find the effect only holds for new users (those whose signup period is shorter than sample median, 22 months, in March 2017).

[4] We do not define "beginning" as January (the starting point of the sample data) or February of 2017 to avoid the impact of seasonal effect from migrants' massive Spring Festival hometown-returning effect.

To summarize, our paper contributes to the literature in the following ways. First, our work contributes to the income inequality literature by providing investigating an important between-group income inequality, supplementing this strand of literature where most existing studies focus on overall income inequality (eg: Demirgüç-Kunt and Levine, 2009; Beck et al., 2007). Second, we add to the economic literature on migration and migrants (eg: Ratha, et al., 2011) where studies investigating the impact from financial market on migrants is scarce. Finally, we add to the Fintech microfinance literature (eg: Jack and Suri, 2014, 2016; Lee, et al., 2017) by focusing Fintech's impact on migrants' livelihood in destination cities while existing literature mainly look at Fintech's influence on migrants' hometown (typically rural area).

2. Literature Review and Contributions

This section discusses in detail four strands of literature related to our paper and our contributions. The literature includes that on traditional non-Fintech based microfinance, Fintech, migration, and income equality.

Microfinance is financial services targeting individuals and small businesses that lack access to conventional banking and related services. Microfinance can potentially reduce income inequality between those that have access to conventional financial services and those that do not. Empirical evidence on this income effect is mixed. Supporting evidence is found in cross-country studies where countries with higher participation of microfinance witness lower income inequality (Hermes 2014; Lacalle-Calderon et al. 2019). However, micro-level studies using randomized controlled trials (RCT) in a few countries fail to find a positive impact of microcredit participation on household income (Angelucci et al, 2015; Attanasio et al., 2015; Augsburg et al, 2015; Crépon et al., 2015; Tarozzi et al, 2015;

Banerjee et al., 2015a, 2015b). Due to the local rather than general effect, Morduch (2020) questions the validity of generalizing the results from the above RCT studies.

Recent years have seen a growing body of literature on Fintech, specifically mobile money, digital loans, and digital economy. Mobile money allows monetary value to be stored on a mobile phone and sent to other users via text messages (Suri and Jack, 2016) and is an important enabler for digital loans (Björkegren and Grissen, 2018). The literature shows that the introduction of mobile money has important welfare effects. For example, mobile money improves risk-coping (Jack and Suri, 2014) and reduces poverty (Suri and Jack, 2016) in Kenya. As to digital credit, the dominant form is consumer credit products, which are typically of short term and high interest rates (Francis et al., 2017). Bharadwaj et al. (2019) find that digital loans raise household resilience to financial shocks. On the other hand, the high interest rate on digital loans may cause difficulty in paying bills (Melzer, 2011) and over-indebtedness and bankruptcy (Skiba and Tobacman, 2019). In addition to consumer credit products, business loans in digital forms are growing recently. Hau, et al. (2019) provide the evidence that compared to traditional bank loans, Fintech credit expands the extensive margin of credit to borrowers of lower credit scores, and thus benefits more people. The theory in Hau, et al. (2019) also predicts that the share of borrowers using the digital credit decreases in the credit scores, implying that digital loans particularly serve low-credit borrowers. Frost, et al. (2019) investigate the development of BigTech credit in Argentina and find that compared to the traditional credit rating service, credit scoring techniques based on big data and machine learning provide a better prediction of loss rates of small businesses.[5] Chen, et al. (2018) use Alibaba's data and find

[5] We use the terms "Fintech", "BigTech" and "TechFin" interchangeably because they all refer to using data generated from online platforms or mobile apps to improve financial services.

that credit access significantly reduces the volatility in firm sales. In sum, we add to this literature by investigating differential impacts of Fintech-based business microloans on migrants versus natives.

Although the literature on migration is large, not much focuses on how financial markets impact migration and the welfare of migrants. Datta (2009) investigates financial exclusions faced by low-paid migrant workers. Jack and Suri (2016) list the reduction of transaction costs of long-distance urban-to-rural remittances as one of the potential channels for lowering extreme poverty and expanding migration to higher-return labor markets. Lee, et al. (2017) find that when mobile banking is introduced to poor rural households and their family members who migrate to the cities in Bangladesh, rural consumption increases and extreme poverty falls. In our paper, we investigate whether urban migrants benefit from the development of Fintech, through the channels other than lower remittance transaction cost. Specifically, we contribute to the literature by examining the financial constraint channel and providing evidence on Fintech's role in narrowing the migrant-native income gap.

Lastly, our paper contributes to the literature on the effect of finance on inequality. Demirgüç-Kunt and Levine (2009) distinguish between the effect of finance on the extensive margin and that on the intensive margin. The extensive margin is to extend the financial services to the individuals who would otherwise have no access to those services. The extensive margin thus benefits the disadvantaged group and decreases the inequality (Beck et al., 2007). The intensive margin is to improve the quality and the range of financial services enjoyed by those already purchasing financial services (Greenwood and Jovanovic, 1990). This tends to widen the inequality. Our research focuses on the extensive margin and the result suggests that Fintech, acting as a more inclusive form of finance to migrants, reduce income inequality across migrants and natives.

3. Fintech Microloans

3.1 Institutional Background

This section provides institutional background on Ant Group, our data provider, and its microloan products. Ant Group is known as the digital finance arm of Alibaba, and its best-known product, Alipay, is the leading third-party payment system in China, with a 55% of the total market share. Our data come from the transaction records of Alipay users' accounts. The Fintech business loans considered in the paper are provided by Ant Group's affiliate, MyBank, an online bank that focuses on small business loans and is rapidly expanding since its incorporation in June 2015. The average loan balance per borrower doubled from RMB 15,000 in 2016 to RMB 31,000 in 2019 and the loan products supported over 20 million small and micro businesses as of 2019, up from 2.77 million borrowers in 2016. The non-performing loan ratio (the amount of loans in default to the total amount of outstanding loans a bank holds) to small and medium-sized businesses has consistently been at around 1%, significantly lower than the industry average of 3.22% in 2019.[6] Given Alipay's leading market share and MyBank's broad coverage of borrowers, the sample data we obtain from Ant Group can reasonably represent the FinTech loan market in China.

---

[6] Data is from ChinaFintech – 2017 Results of Ant's MyBank (https://chinatechecon.wordpress.com/tag/zhejiang-e-commerce-bank/), and MYbank Served Over 20 Million SMEs as of 2019, Further Spurring the Growth of China's Small and Micro Businesses (https://www.businesswire.com/news/home/20200427005353/en/MYbank-Served-20-Million-SMEs-2019-Spurring).

The business microloans offered by MyBank have the features of a typical Fintech credit product: instant, automated and remote (Chen and Mazer, 2016). First, it is almost instantaneous from loan application through approval, with a 3-minute application and nearly 0-second approval. Second, the evaluation of loan applications is automated with big data technology accomplishing the screening instead of relying on human judgement. Third, the entire process of loan application, approval, and payments is done online without the need of visiting a branch or agent in person.

3.2 Traditional Banking vs. Fintech Loans

Existing theories suggest that financial market imperfections prevent poor people from escaping poverty by limiting their access to formal financial services, and thus new financial technologies become important tools to enable financial inclusion (Demir et al., 2020; Galor and Zeira, 1993).

The information asymmetry problem faced by traditional financial intermediaries leads to financial exclusion for smaller businesses and unprivileged entrepreneurs. Traditional banking relies on three main means to reduce information asymmetry: financial statements, collateral, and soft information collection (Figure 1 Panel A). First, financial data provided by smaller businesses are often less credible, and may not have enough profit years that would warrant a loan. Second, entrepreneurs with more endowment, such as real estate, vehicles, and parental wealth, can more easily provide collaterals for loans. Migrant entrepreneurs of small businesses, in contrast, are less able to do so because they are often born in a less wealthy family or face restrictions in buying a house in cities that they are not originally from. Third, soft information is collected through frequent personal contacts between the borrower and the loan officer. This process of relationship building requires time and effort, and borrowers from small businesses lack the resources (e.g., connection, special manpower in public relations) to engage in the relationship building. Taken together, in the traditional banking system, financial resources are more likely to be allocated to larger firms and financially sound entrepreneurs, thereby widening the inequality in economic opportunities.

[Figure 1 here]

Digital finance, nevertheless, resolves the above information problem and thus offers more equal economic opportunities than the traditional banking system (Bharadwaj, et al., 2019). Specifically, as shown in Figure 1 Panel B, besides financial statements, Fintech firms have a large dataset of their users' daily transactions, which provide current and detailed information on a business' operating performance and behavioral attributes on their willingness and ability to repay a loan. As a result, collateral often becomes unnecessary in the Fintech loans due to the availability of additional information generated from big data, allowing these loans more accessible to entrepreneurs without much endowment. In addition, with Fintech, marginal cost of dispersing a Fintech loan to an additional user is near zero (Bjorkegren and Grissen 2018). This makes smaller and cheaper loans in a much larger scale than any traditional financial intermediaries can do. Furthermore, Fintech clients' hard information is automatically generated on a daily basis and their soft information is hardened through being handled by models and machines rather than humans. Hence, under the Fintech platform, all businesses and entrepreneurs are treated equally and smaller businesses are no longer disadvantaged. Overall, Fintech improves data quality and lowers data collection cost, thus reducing the information costs for smaller businesses and migrant entrepreneurs who otherwise have asymmetric information problems. Digital finance, as a result, offers more equal economic opportunities, especially for previously unprivileged groups.

4. Data and Main Variables

Our data is provided by Ant Group and covers 167,268 active individual business owners from January 2017 to July 2019. The data is anonymous and desensitized to preserve confidentiality. This dataset provides a rich array of variables, including each individual business owner's monthly operating performance, indicator variables on whether the business owners draw business loans each month, whether a business owner is a migrant or a native, and whether a business owner conducts the business online on the Taobao platform or offline using a QR code to receive payments. Other individual-level information includes the business owners' demographics, the industry that the business is in, resident city, etc. City level variables include the level of city development and various indexes measuring a city's mobile payment penetration.

Our paper focuses on self-employed workers and small business owners whose business revenue is likely to compose the majority of their personal income. As a result, we require that our sample includes only the individuals who 1) register their businesses using their personal identification information instead of corporate information; and 2) are active users (i.e., having at least 6-month consecutive transaction records in Alipay database by the end of 2018).[7] Ant Group provides us with a random draw of about 167 thousand users based on the criteria above, and this random sample includes around 4 million user-month observations.

[7] Non-active users are likely to be consumers rather than business owners who use their Alipay accounts for business operations. Anyone can apply for a personal QR code and use it for money transfer among friends and relatives. These non-commercial transactions probably happen only occasionally and hence the criteria we impose can filter out non-business owner users.

The main dependent variable is the logarithm of monthly revenue of individual business owners. This is the sum of the money inflow into an Alipay user's account from transactions (not from loans) each month. Although these micro-business owners may have income sources other than their owned micro-business, the difference of the above revenue variable between migrant and native business owners could imply whether the income gap widens or narrows. Two other performance measures, the number of transactions and the number of unique customers each month, can also help us gauge how well the business is performing. Due to the right skewness, all the three operating performance variables are winsorized at the 99% level.

To compare migrants and natives, we define an indicator variable, which is equal to one if the business owner is a migrant and zero otherwise. We identify an individual as a migrant if he/she resides in a province different from the one where he/she was born. We then define an indicator variable equal to one if the business owner draws a business loan during the month, and zero otherwise.[8] We provide the detailed definitions of all the variables in Online Appendix Table A1.

Table 1 presents the summary statistics of the main variables included in the dataset. The monthly revenue (*revenue*), has a median of ￥5,253.00 (US$763.68), much smaller than its mean ￥27,329.35 (US$3,973.16), indicating right skewness. The small dollar value of the revenue confirms that the typical sample business owner is a micro-business owner. The average *Migrant* indicator, 0.27, shows that three out of ten business owners are migrants. One out of four business owners operates online (with the online seller

[8] Due to confidentiality reason, the data does not allow us to obtain details about the loan (amount, interest rate, duration, covenants, etc.).

indicator, *Taobao_seller*, being equal to one) and 88% business owners accept QR code payments (QR-code business owner indicator, *IsQRcode*, is one). The prevalence of QR-code business owners is not surprising due to the low cost and low technology barrier of QR codes. Even very small street-vendors who do not own a computer or do not know anything about webpage design can easily apply for a QR code with a mobile phone and use it to receive payments for their businesses. The business owners in the sample have a median age of 34. The median gender is 1, meaning that the typical sample business owner is male. The medians of *Child_score*, *RealEstate* and *Marriage* are 0.68, 0.57 and 0.66, suggesting that the sample business owners' probabilities of having at least a child, having real property, and being married are more than half. The average *DrawLoan* is 0.13. That is, 13% of the monthly sample data are associated with drawing loans, meaning that a representative business owner draws microloans once or twice per year.

[Table 1 here]

5. Empirical Results

5.1. Effect of Fintech Microloans on Income: Migrant vs. Native Business Owners

5.1.1. Empirical Model

To understand the effect of the Fintech microloan on migrant business owners' income compared to their native counterparts, we employ a standard DID setting:

$$\ln(revenue_{i,t}) = b_1 Migrant_i + b_2 DrawLoan_{i,t} + b_3 Migrant_i \times DrawLoan_{i,t} + Ctrls + A_c + B_t + \varepsilon_{i,t} \quad (1),$$

where $revenue_{it}$ is the revenue received from business transactions by business owner i in month t, $Migrant_i$ is an indicator variable on whether business owner i is identified as a migrant who works in a different province from where she was born, and $DrawLoan_{it}$ is an indicator variable showing whether business owner i draws loan credit in month t. $Ctrls$ refer to the control variables, including city development degree dummies, city gdp per capita in logarithm and a set of user characteristics. These characteristics include online-business owner dummy, QR-code business owner dummy, farm-loan user dummy, gender, age, squared age, probability of being married, probability of having a baby, probability of owning real estate, non-missing education dummy, occupation dummies, industry dummies and number of months since first signing up with Alipay. Variable definitions are in Online Appendix Table A1. A_c and B_t refer to city fixed effects and month fixed effects.

The coefficient of interest is b_3. If Fintech microloans benefit migrants more than natives, we expect to see b_3 being positive and significant.

5.1.2. Matching

To provide a comparable basis for the analysis, the regression is run on a matched sample of migrant and native business owners who are similar in observable dimensions except for whether they are migrants. This is to remove any potential pre-treatment differences between migrants and natives.

The detailed steps of constructing the matching sample are as follows. First, we use a matching period of March to May of 2017 (we do not use January or February of 2017 for matching to avoid the impact of seasonal effect from migrants' massive Spring Festival hometown-returning effect.) and remove the business owners who register with Alipay but have not been operating by keeping only business owners who have monthly revenues greater than ￥1.00 ($0.15) for all the 31 months in our sample. We also restrict our sample in cities of higher development level identified by cities with the first three degrees where nearly 90% of the migrant business owners live in, as migrants who go to larger cities (as opposed to small ones) are of our interest. Only individuals who do not draw loans during the matching period participate in the matching process so that we can have a sample where migrants and natives are similar when FinTech loan hasn't come in.

Second, for each migrant, we look for one similar native match by nearest neighbor matching). Specifically, for each migrant, we find candidate matches of native observation in the same city, same month, same industry,[9] same education level, similar age (difference in age is less or equal than 30), and similar performance (that is, the natives who see revenue in 2017M3 within ±10% range of the migrants' revenue in the same month, one-month growth rate within ±30% that of the migrants, and two-month growth rate within ±40% that of the migrants and the difference in number of customers is less than the amount of one standard deviation for *Ncustomers*) We do this to ensure the migrant business owner and their native counterparts are comparable ex ante. Then, among these native candidates, we keep one closest native match based on the Euclidean distance of the migrant's and the native's one-month revenue growth rate during 2017M3-2017M5.

Third, we link these matched observations to the regression variables, and that constructs our main matched sample dataset.

[9] To raise the possibility of getting a match, based on the industry information provided in the dataset, we only do a rough 4-category industry classification: retail, catering, other services and the rest of industries, which accounts for 40.6%, 32.5%, 10.6% and 16.4% of the observations in the full sample, respectively.

We do a balance test to investigate whether migrants and their matched natives are similar in all observable dimensions. Panel A of Table 2 shows the statistics for migrants and natives on 14 characteristics including performance, business owner type and demographics. The difference is insignificant at 5% level for all variables. Only two variables, revenue of 2017M3 and revenue growth from 2017M3 to 2017M4, are significantly different at 10% level between migrants and natives. However, the magnitude of the difference is not large: the revenue difference for 2017M3 is ￥385.52 ($56.05) and the one-month revenue growth difference is -1.01%. We conclude that the migrants and natives in the matched sample are similar at the beginning periods of our full sample.

To see whether the matched sample is similar to the full sample, Panel B of Table 2 presents a sample representative analysis where we compare summary statistics for the full sample and the matched sample for the 13 variables used in the regressions. The statistics are similar on demographic variables, but different on business-related variables. Specifically, matched sample businesses in general performs better than those in full sample (eg: p25 of matched sample *Revenue* is larger than p50 of full sample), signed up in the platform earlier, and see a much larger portion operating online and less portion using QR money-receiving code (that is, less likely to be offline). Such difference is not surprising as we restrict our sample to those who operate continuously for all 31 months in the sample before matching, while full sample includes many other businesses that started or stopped operating during the sample period.

[Table 2 here]

5.1.3. Regression Results

We conduct the DID regression in Equation (1) using the matched sample and report theresult in Table 3. The coefficient of the interaction term, $Migrant \times DrawLoan$, is positive and significant, indicating migrant business owners harvest more revenue from using the Fintech microloans compared to the native owners. In Column (2), we include only business owners who draw loans at least once during the sample period, so we ensure that the included entrepreneurs can and are willing to use Fintech microloans. The direction and statistical significance of the coefficients of interest remain unchanged. In terms of the magnitude, the coefficients indicate that Fintech microloans give migrants an additional 14.06% (full-sample estimate in Column 1) or 8.89% (subsample estimate that requires each business owner to have drawn loans in Column 2) boost in revenue compared to their native matches. Translated into dollar amount based on sample revenue median (￥34,457 or $5009.38), the additional revenue gain for migrants relative to natives equals $704.32 (full-sample estimate) or $445.33 (sub-sample estimate). The coefficient of $DrawLoan$ is also positive, indicating that business owners who take out a loan generate a greater income than those who do not.

In terms of the control variables, we find that Taobao online shop owners show higher revenue and QR-code business owners with lower revenue. Opening an online shop is more complicated than registering for a money-receiving QR code, so it’s not surprising that online business owners tend to be more sophisticated and earn more. It also concurs with anecdotal evidence that QR-code business owners tend to be much smaller (such as street vendors). We also find that those who are female, have signed up with Alipay earlier, and less likely to be married earn higher revenues.

[Table 3 here]

Table 4 provides a robustness check for the main regression results. We replace the dependent variable, revenue, with alternative business activity measures, the logarithm of number of transactions and the number of unique customers. Consistent with the main regression, we find positive and significant coefficients for both *DrawLoan* and the interaction term, *Migrant×DrawLoan*, and for both the full sample and the subsample that requires all business owners to have drawn loans at least once.

[Table 4 here]

Overall, Tables 3 and 4 provides evidence that migrants benefit more from Fintech microloans than natives do.

5.2. Heterogeneity Analysis

Having established the main result that migrants generate more revenues than natives after using Fintech loans, we now examine the cross-sectional differences in the main result.

5.2.1 Financial Constraint

First, we explore whether financial constraint affects the income converging effect. Existing theory shows that finance services enhance equality only when the service is provided to the "extensive margin"; that is, the individuals who would otherwise have no access to those services (Beck et al., 2007). We thus conjecture that the income-enhancing effect should be more prominent for migrants who are more financially constrained and have less access to alternative financial services.

We test this financial constraint hypothesis by adding additional interaction terms:

$$\ln\,(revenue_{i,t}) = \gamma_0 FinConstraint_i + (\delta_1 + \gamma_1 FinConstraint_i) \cdot Migrant_i + (\delta_2 + \gamma_2 FinConstraint_i) \cdot DrawLoan_{i,t} + (\delta_3 + \gamma_3 FinConstraint_i) \cdot Migrant_i \times DrawLoan_{i,t} + Ctrls + A_c + B_t + \varepsilon_{i,t}\ , \qquad (2)$$

where *FinConstraint* is the indicator for individuals with more financial constraints. The definitions of other variables are the same as in Equation (1).

We use two ways to identify individuals with more financial constraint. The first one is low initial revenue. These individuals are most likely to have lower income and face the risk of being financially excluded. Specifically, we identify individuals with more financial constraint by revenue of 2017M3 below median, below 25th percentile or within the first tercile. The second way is to find businesses that only have a QR money-receiving code and do not operate online. Online businesses require more resources (e.g., laptop, network) and be more sophisticated (e.g., webpage design) than an offline business receiving payment with a QR code. There is also anecdotal evidence that many small street vendors who used to receive payment solely by cash seldom had access to traditional loans. Many of them now also receive payments by QR code. These offline businesses owning a QR code are likely to be more financially constrained and thus we use this as an identification.

If the financial constraint hypothesis is valid, we shall see stronger results for migrants with low probability of owning a property. The result in Table 5 confirms this conjecture. The coefficient, γ_3, is positive and significant for both identifications of *FinConstraint*, indicating that the additional revenue gain for migrants over natives following using a Fintech credit is more prominent for those migrants who are more financially constrained. This is in line with our conjecture that Fintech microloan helps migrants by making financial resources more accessible, alleviating their financial constraint. To sum up, the results here suggest a positive role of microloans in supporting the migrants by relaxing their financial constraints.

[Table 5 here]

5.2.2 Economic Development Level

We then investigate how economic development affects the impact of Fintech on migrant-native income gap. We use two measures of economic development. In the first measure, we sort all cities in the matched sample based on their GDP per capita in year 2017 and classify them into low, medium and high economic development levels based on terciles. The second measure is based on geographic regions. Southern and eastern China achieve greater economic success than the rest of the country due to historical reasons (that is, southern and eastern regions experiment economic reforms earlier than other regions). We compare south (more developed) vs. north (less developed), and east (more developed) vs. non-east (less developed). Specifically, we use Qinling Mountains-Huaihe River to divide China into the North and South regions and follow National Bureau of Statistics of China to define East area (more developed) and combine the rest of the areas as non-East area (less developed).

Table 6 presents the result with different economic development levels. We find that the positive interaction term is only present in cities or areas with more advanced economic development: cities with high GDP per capita, southern area, and eastern area. In other areas, the interaction term is either insignificant or negative with a positive significant coefficient for *Migrant*. That is, migrants experience incremental gain more than natives do only in better developed regions. A possible explanation is that migrants are more likely to be unprivileged compared to natives in more developed areas, and thus are likely to have more additional gain than natives after taking Fintech loans.

[Table 6 here]

5.2.3 Risk Attitude

We explore potential heterogeneity among groups with different risk attitudes: does the income converging effect we find occurs in less or more risk-averse migrants?

We use gender and age of the business owners as proxies of risk attitudes. Previous finance and psychology literature finds that men are overconfident relative to women, and women tend to be less risk-taking in executive decisions (eg: Huang and Kisgen, 2013). Previous studies also show the evidence that risk-taking decreases with age (e.g., Jianakoplos and Bernasek, 2006). We thus split the sample by the sample median of individuals' ages, with the younger group (whose age is lower than sample median) considered as more risk-taking.

Table 7 presents the result based on groups with different risk attitudes. The interaction term, *Migrant×DrawLoan*, is positive and significant only in the female group and the elder group, suggesting that the income converging effect of Fintech loans is only present in risk-averse business owners. One possible explanation is that risk-taking groups may have already explored other financing sources (such as P2P), while risk-averse individuals have tighter resources, and thus using Fintech loans is associated with more gains for the latter group.

[Table 7 here]

5.2.4 Signup Period

We then examine how the length of the business owner's Fintech signup period affects the income converging effect. We define *sign up early* group as those whose signup period is longer than sample median (22 months) in March 2017, and *sign up late* group as those whose signup period is shorter than sample median.

Table 8 shows that the positive interaction term only occurs in the *sign up late* group, suggesting that migrants who are relatively new to the Fintech platform gain more than natives after using the Fintech loans. According to the firm life cycle theory (eg: Bulan and Subramanian, 2009), a business generally follows a life-cycle trajectory from origin to maturity, associated with a shrinking investment opportunity and decreasing cost of raising external capital. Thus, external financing resources such as Fintech loans are more relevant for new businesses. Another possibility is associated with the way businesses are operated. Among *sign up late* group, we have almost equal number of businesses that have QR money-receiving code and that owns online Taobao shop (502 vs. 499), while among *sign up early* group, over 92% of all the businesses have an online shop. Thus the insignificant interaction term for *sign up early* group could be associated with online business operation because whether the business owner lives in his hometown (as native) or another province (as migrant) becomes less relevant when the main business is operated online.

[Table 8 here]

6. Conclusion

This paper examines the effect of Fintech-based microfinance on the revenue of migrant versus native business owners who run personal business using Alipay's business owner level data. Results show that microfinance brings incremental revenue increase for migrants compared to their native counterparts. Although these micro-business owners may have income sources other than their owned micro-business, the difference of the revenue variable between migrant and native business owners could imply whether the income gap widens or narrows. The overall evidence suggests Fintech-based microfinance narrows the income inequality between the migrant and the native.

More detailed analyses show that such the narrowed income gap between the migrant and the native mainly exists in the businesses with more financial constraint, in more economically developed areas, for more risk averse business owners, and among the owners who are relatively new to the Fintech platform. .

For policy makers who intend to attract migrants into the city or to enhance the economic integration of migrant population into native society, being open to the penetration of Fintech loans would be of help.

References

Albareto, G., & Mistrulli, P. E. (2011). Bridging the gap between migrants and the banking system. *Bank of Italy Temi di Discussione (Working Paper) No*, *794*.

Anderloni, L. and Vandone, D., 2006. New market segments: migrants and financial innovation. *UNIMI-Economics, Business, and Statistics Research Working Paper*, (2).

Angelucci, M., Karlan, D. and Zinman, J., 2015. Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco. *American Economic Journal: Applied Economics*, *7*(1), pp.151-82.

Augsburg, B., De Haas, R., Harmgart, H. and Meghir, C., 2015. The impacts of microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics*, *7*(1), pp.183-203.

Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E. and Harmgart, H., 2015. The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *American Economic Journal: Applied Economics*, *7*(1), pp.90-122.

Banerjee, A., Duflo, E., Glennerster, R. and Kinnan, C., 2015. The miracle of microfinance? Evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, *7*(1), pp.22-53.

Banerjee, A., Karlan, D. and Zinman, J., 2015. Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, *7*(1), pp.1-21.

Banerjee, A., Duflo, E. and Hornbeck, R., 2018. How much do existing borrowers value Microfinance? Evidence from an experiment on bundling Microcredit and insurance. *Economica*, *85*(340), pp.671-700.

Beck, T., Demirgüç-Kunt, A. and Levine, R., 2007. Finance, inequality and the poor. *Journal of Economic Growth*, *12*(1), pp.27-49.

Bharadwaj, P., Jack, W. and Suri, T., 2019. Fintech and household resilience to shocks: Evidence from digital loans in Kenya (No. w25604). *NBER working paper*.

Björkegren, D. and Grissen, D., 2018, May. The potential of digital credit to bank the poor. In *AEA Papers and Proceedings* (Vol. 108, pp. 68-71).

Bulan, L.T. and Subramanian, N., 2009. The firm life cycle theory of dividends. *Dividends and dividend policy*, *1*, pp.201-213.

Chen, T., Huang, Y., Chen, L. and Sheng, Z., 2018. Finance and firm volatility (No. BOOK). [sn].

Chen, G. and Mazer, R., 2016. Instant, automated, remote: The key attributes of digital credit. *CGAP Blog*, *February*, *8*, p.2016.

Cingano, F., 2014. Trends in income inequality and its impact on economic growth. *OECD Social, Employment and Migration Working Papers*, No. 163, OECD Publishing, Paris, https://doi.org/10.1787/5jxrjncwxv6j-en.

Crépon, B., Devoto, F., Duflo, E. and Parienté, W., 2015. Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *American Economic Journal: Applied Economics*, *7*(1), pp.123-50.

Datta, K., 2009. Risky migrants? Low-paid migrant workers coping with financial exclusion in London. *European Urban and Regional Studies*, *16*(4), pp.331-344.

Demir, A., Pesqué-Cela, V., Altunbas, Y. and Murinde, V., 2020. Fintech, financial inclusion and income inequality: a quantile regression approach. *The European Journal of Finance*, pp.1-22.

Demirgüç-Kunt, A. and Levine, R., 2009. Finance and inequality: Theory and evidence. *Annu. Rev. Financ. Econ.*, *1*(1), pp.287-318.

Francis, E., Blumenstock, J. and Robinson, J., 2017. Digital credit in emerging markets: A snapshot of the current landscape and open research questions. *Journal of Digital Commerce*, *7*(19), pp.345-351.

Friedman, E. and Lee, C.K., 2010. Remaking the world of Chinese labour: A 30-year retrospective. *British journal of industrial relations*, *48*(3), pp.507-533.

Frost, J., Gambacorta, L., Huang, Y., Shin, H.S. and Zbinden, P., 2019. BigTech and the changing structure of financial intermediation. *Economic Policy*.

Galor, O. and Zeira, J., 1993. Income distribution and macroeconomics. *The review of economic studies*, *60*(1), pp.35-52.

Greenwood, J. and Jovanovic, B., 1990. Financial development, growth, and the distribution of income. *Journal of Political Economy*, *98*(5, Part 1), pp.1076-1107.

Hao, T., Sun, R., Tombe, T. and Zhu, X., 2020. The effect of migration policy on growth, structural change, and regional inequality in China. *Journal of Monetary Economics*.

Hau, H., Huang, Y., Shan, H. and Sheng, Z., 2019. How FinTech enters China's credit market. *AEA Papers and Proceedings* Vol. 109, pp. 60-64.

Hermes, N., 2014. Does microfinance affect income inequality?. *Applied Economics*, *46*(9), pp.1021-1034.

Huang, J. and Kisgen, D.J., 2013. Gender and corporate finance: Are male executives overconfident relative to female executives?. *Journal of financial Economics*, *108*(3), pp.822-839.

Jack, W. and Suri, T., 2014. Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, *104*(1), pp.183-223.

Jianakoplos, N.A. and Bernasek, A., 2006. Financial risk taking by age and birth cohort. *Southern Economic Journal*, pp.981-1001.

Kaboski, J.P. and Townsend, R.M., 2012. The impact of credit on village economies. *American Economic Journal: Applied Economics*, *4*(2), pp.98-133.

Karlan, D. and Zinman, J., 2010. Expanding credit access: Using randomized supply decisions to estimate the impacts. *The Review of Financial Studies*, *23*(1), pp.433-464.

Knight, J., Deng, Q. and Li, S., 2011. The puzzle of migrant labour shortage and rural labour surplus in China. *China Economic Review*, *22*(4), pp.585-600.

Lacalle-Calderon, M., Larrú, J.M., Garrido, S.R. and Perez-Trujillo, M., 2019. Microfinance and income inequality: New macrolevel evidence. *Review of Development Economics*, *23*(2), pp.860-876.

Lee, J.N., Morduch, J., Ravindran, S., Shonchoy, A. and Zaman, H., 2017. Poverty and migration in the digital age: Experimental evidence on mobile banking in Bangladesh. *American Economic Journal: Applied Economics*.

Li, S.M., 2010. Mortgage loan as a means of home finance in urban China: A comparative study of Guangzhou and Shanghai. *Housing Studies*, *25*(6), pp.857-876.

Melzer, B.T., 2011. The real costs of credit access: Evidence from the payday lending market. *The Quarterly Journal of Economics*, *126*(1), pp.517-555.

Meng, X. and Bai, N., 2007. How much have the wages of unskilled workers in China increased? *China: Linking markets for growth*, pp.151-175.

Morduch, J., 2020. Why RCTs failed to answer the biggest questions about microcredit impact. *World Development*, *127*, p.104818.

Ratha, D., Mohapatra, S., & Scheja, E. (2011). Impact of migration on economic and social development: A review of evidence and emerging issues. World Bank Policy Research Working Paper 5558.

Pitt, M.M. and Khandker, S.R., 1998. The impact of group-based credit programs on poor households in Bangladesh: Does the gender of participants matter? *Journal of Political Economy*, *106*(5), pp.958-996.

Skiba, P.M. and Tobacman, J., 2019. Do payday loans cause bankruptcy? *The Journal of Law and Economics*, *62*(3), pp.485-519.

Suri, T. and Jack, W., 2016. The long-run poverty and gender impacts of mobile money. *Science*, *354*(6317), pp.1288-1292.

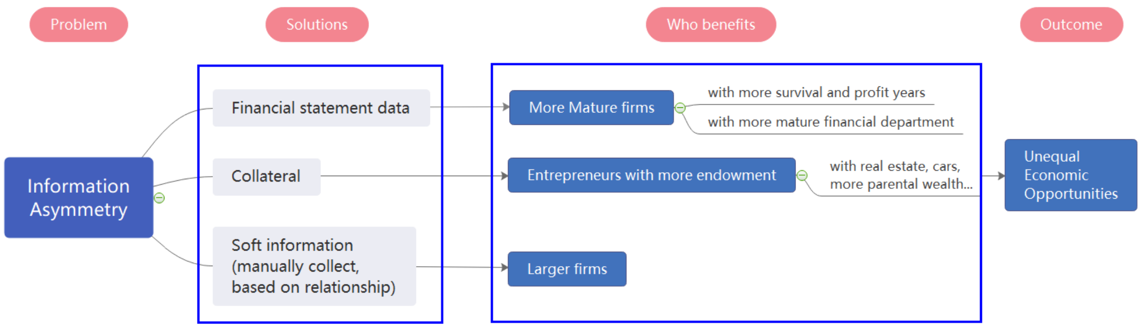
Tarozzi, A., Desai, J. and Johnson, K., 2015. The impacts of microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics*, *7*(1), pp.54-89.

Wang, J.J. and Tian, Q., 2014. Consumer vulnerability and marketplace exclusion: A case of rural migrants and financial services in China. *Journal of Macromarketing*, *34*(1), pp.45-56.

Figure 1: Finance and Inequality: Traditional Banking

The figure shows how the approaches in which banking deals with information asymmetry between lenders and borrowers lead to unequal economic opportunities for larger versus smaller businesses. Panel A is the case of traditional banking and Panel B is the case of digital finance.

Panel A: Traditional Banking



Panel B: Fintech Credit

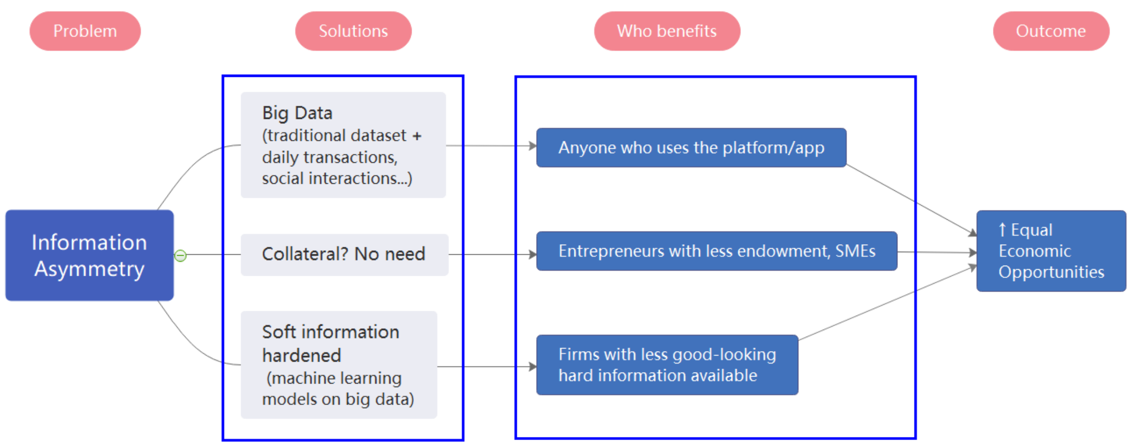


Table 1: Summary Statistics of Main Variables for Individual Business owners

This table reports the descriptive statistics for micro-level individual business owner data. Revenues, number of transactions (*Ntransactions*), and number of customers (*Ncustomers*) are the data for each business owner each month. All dollar amounts are converted from Chinese Yuan to US dollars based on the exchange rate on December 31, 2018. Variable definitions are in Online Appendix Table A1.

Variables	**# Obs.**	**Mean**	**SD**	**p25**	**Median**	**p75**
Dependent variable: Main						
Revenue	4,181,955	￥27,329.35 ($3,973.16)	￥71,071.62 ($10,332.43)	￥1,425.50 ($207.24)	￥5,253.00 ($763.68)	￥18,406.00 ($2,675.87)
Dependent variable: Alternative						
Ntransactions	4,181,955	236.26	528.37	24	68	197
Ncustomers	4,181,955	164.07	391.87	13	46	131
Independent variables:						
DrawLoan	4,181,955	0.13	0.34	0	0	0
Migrant	4,181,955	0.27	0.44	0	0	1
Control variables:						
Taobao_seller	4,181,955	0.25	0.43	0	0	0
IsQRcode	4,181,955	0.88	0.32	0	1	1
Age	4,181,955	36.14	9.20	29	34	42
Gender	4,181,955	0.5780	0.4939	0	1	1
Signup_period (months)	4,181,955	18.85	16.64	7	15	24
Child_score	3,978,471	0.6037	0.3473	0.2595	0.6797	0.9541
RealEstate	4,178,807	0.6089	0.1712	0.4928	0.5662	0.6665
Marriage	3,978,471	0.5963	0.2638	0.3945	0.6588	0.8181
City_degree	4,181,955	2.83	1.19	2	3	4
Gdppc (city level)	4,176,370	￥84,417 ($12,273)	￥45,665 ($6,639)	￥44,672 ($6,494)	￥75,987 ($11,047)	￥118,015 ($17,157)

Table 2: Summary Statistics and Sample Representativeness of the Matched Sample

This table reports descriptive statistics for matched sample and its sample representativeness compared to the full sample. Panel A shows summary statistics for migrants and matched natives during matching period, 2017M3-2017M5. For each variable, we provide the average (Mean) and standard deviation (STD) as well as the t-statistics for difference between migrant group and native group. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. Panel B presents summary statistics of the variables in the full sample and matched sample. For each variable, we provide the average (Mean), standard deviation (STD), 25th percentile (p25), 50th percentile (p50) and 75th percentile (p75). Revenues, number of transactions (*Ntransactions*), and number of customers (*Ncustomers*) are the data for each business owner each month. Variable definitions are in Online Appendix Table A1.

Panel A: Summary Statistics for Migrants and Matched Natives during Matching Period

	Matched Sample: Migrants		Matched Sample: Natives		*t*-statistic for difference
	(1)	(2)	(3)	(4)	(5)
Variable	Mean	STD	Mean	STD	(1)-(3)
$Revenue_{2017M5}$	￥59605.62	77910.37	￥59695.73	76841.24	-0.1322
$Revenue_{2017M4}$	￥62342.03	80282.19	￥62945.48	80748.74	-1.0539
$Revenue_{2017M3}$	￥72062.87	89795.83	￥71677.35	89523.33	1.7871*
$Ncustomers_{2017M5}$	380.59	544.69	390.74	541.83	-0.6675
$Ntransactions_{2017M5}$	474.29	698.24	495.18	710.98	-1.0313
$g\_Revenue_{2017M3,2017M4}$	-0.1017	0.2799	-0.0916	0.2707	-1.9253*
$g\_Revenue_{2017M4,2017M5}$	-0.0136	0.2897	-0.0205	0.2890	1.3261
$g\_Revenue_{2017M3,2017M5}$	-0.1066	0.3921	-0.1024	0.3883	-0.6533
Taobao_seller	0.8049	0.3965	0.7902	0.4074	0.9096
IsQRcode	0.6000	0.4902	0.5744	0.4947	1.2068
Farm_loan	0.2293	0.4206	0.2305	0.4214	-0.0940
Gender	0.5951	0.4912	0.5707	0.4953	1.0370
Age	35.7024	8.4916	36.0512	9.0649	-0.8623
Signup_period	30.7415	17.2603	30.7037	17.5419	0.0464

Panel B: Sample Representative Analysis

Variable	Mean	STD	p25	p50	p75
Revenue					
Full Sample	￥ 27,329	￥ 71,072	￥ 1,426	￥ 5,253	￥ 18,406
Matched Sample	￥ 70,806	￥ 97,564	￥ 13,909	￥ 34,457	￥ 82,200

Ntransactions					
Full Sample	236.2578	528.3743	24	68	197
Matched Sample	509.1268	772.3509	79	208	564
Ncustomers					
Full Sample	164.073	391.872	13	46	131
Matched Sample	396.6215	596.9597	63	164	445
Taobao_seller					
Full Sample	0.2492	0.4325	0	0	0
Matched Sample	0.7976	0.4018	1	1	1
IsQRcode					
Full Sample	0.8803	0.3246	1	1	1
Matched Sample	0.5872	0.4923	0	1	1
Farm_loan					
Full Sample	0.3615	0.4804	0	0	1
Matched Sample	0.2299	0.4208	0	0	0
Edu_notmissing					
Full Sample	0.0431	0.2032	0	0	0
Matched Sample	0.0024	0.0493	0	0	0
Gender					
Full Sample	0.5780	0.4939	0	1	1
Matched Sample	0.5829	0.4931	0	1	1
Age					
Full Sample	36.1400	9.2039	29	34	42
Matched Sample	35.8768	8.7794	30	34	40
Signup_period					
Full Sample	18.8470	16.6434	7	15	24
Matched Sample	40.7226	19.5565	25	39	56
Child_score					
Full Sample	0.6037	0.3473	0.2595	0.6797	0.9541
Matched Sample	0.6547	0.3294	0.3449	0.7732	0.9684
RealEstate					
Full Sample	0.6089	0.1712	0.4928	0.5662	0.6665
Matched Sample	0.5948	0.1692	0.4888	0.5490	0.6232
Marriage					

Full Sample	0.5963	0.2638	0.3945	0.6588	0.8181
Matched Sample	0.6119	0.2650	0.4414	0.6771	0.8326

Table 3: Effect of Microloans on Business owners' Revenues: Migrants vs. Natives

This table presents the impact of microloans on migrant business owners' revenues compared to their native counterparts. We report the result of the DID regression in Equation (1), using the migrant business owners and the native business owners matched by city, month, industry, and revenues of past three months. $Migrant_i$ is an indicator variable equal to one if business owner i is a migrant (his/her residential province as of the end of Dec. 2018 differs from the province where he/she was born), and zero otherwise. $DrawLoan_{it}$ is an indicator variable equal to one if business owner i draws loan credit during month t. Other variable definitions are in Online Appendix Table A1. Column (1) reports the result with all observations in the matched sample, while Column (2) reports the result on the subsample that contains observations for business owners who take out a microloan at least once in the whole sample period. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Dep. var. = *Ln(Revenue)*	(1) Full sample	(2) including only business owners who draw loans at least once during the sample period
$Migrant_i$	0.0020	0.0467***
	[0.0122]	[0.0152]
$DrawLoan_{it}$	0.1097***	0.1190***
	[0.0254]	[0.0260]
$Migrant_i \times DrawLoan_{it}$	0.1406***	0.0889**
	[0.0339]	[0.0348]
Taobao_seller	0.3736***	-0.0450
	[0.0339]	[0.0625]
IsQRcode	-0.0815***	-0.1334***
	[0.0142]	[0.0171]
Farm_loan	0.0626***	-0.0558*
	[0.0229]	[0.0287]
Ln(gdppc)	-0.0405**	-0.0469*
	[0.0177]	[0.0209]
Gender	-0.0457***	-0.0270**
	[0.0122]	[0.0146]
Age	-0.0046	-0.0112*
	[0.0068]	[0.0090]
Age_sqr	-0.0001	-0.0000
	[0.0001]	[0.0001]
Edu_notmissing	0.0343	0.0911

	[0.0872]	[0.00809]
Child_score	-0.0656***	0.0425***
	[0.0066]	[0.0260]
RealEstate	0.2010***	0.2573
	[0.0345]	[0.0417]
Marriage	-0.5910***	-0.5967***
	[0.0321]	[0.0401]
Signup_period	0.0061***	0.0071***
	[0.0004]	[0.0004]
City_degree1	17.7040***	18.019***
	[0.3775]	[0.4248]
City_degree2	18.371***	18.608***
	[0.3849]	[0.4329]
City FE	Yes	Yes
Industry FE	Yes	Yes
Occupation FE	Yes	Yes
Month FE	Yes	Yes
# Obs.	50,840	35,712
Adj R2	0.0301	0.0340

Table 4: Effect of Microloans on Business owners' Transaction Activities and Customers: Migrants vs. Natives

This table presents the impact of microloans on migrant business owners' transaction activities and customers compared to their native counterparts. We report the result of the DID regression in Equation (1), using the migrant business owners and the native business owners matched by city, month, industry and number of transactions of past three months. The dependent variable is the natural logarithm of the number of recorded transactions each month for Columns (1) and (2), and is the natural logarithm of the number of unique customers each month for Columns (3) and (4). $Migrant_i$ is an indicator variable equal to one if business owner *i* is a migrant (his/her residential province as of the end of Dec. 2018 differs from the province where he/she was born), and zero otherwise. $DrawLoan_{it}$ is an indicator variable equal to one if business owner *i* draws loan credit during month *t*. Other variable definitions are in Online Appendix Table A1. Columns (1) and (3) report the result with all observations in the matched sample, while Columns (2) and (4) reports the result on the subsample that contains observations for business owners who take out a microloan at least once in the whole sample period. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	Dep. var. = *Ln(Ntransactions)*		Dep. var. = *Ln(Ncustomers)*	
	(1) Full sample	(2) including only business owners who draw loans at least once during the sample period	(3) Full sample	(4) including only business owners who draw loans at least once during the sample period
$Migrant_i$	-0.0443***	-0.0521***	-0.0310***	-0.0399**
	[0.0125]	[0.0155]	[0.0124]	[0.0155]
$DrawLoan_{it}$	0.1894***	0.1573***	0.2079***	0.1627***
	[0.0250]	[0.0259]	[0.0251]	[0.0260]
$Migrant_i$	0.0971***	0.1089***	0.1085***	0.1191***
$\times DrawLoan_{it}$	[0.0349]	[0.0360]	[0.0348]	[0.0359]
Controls	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Nobs	50,840	35,712	50,840	35,712
Adj R2	0.0634	0.0661	0.0678	0.0652

Table 5: Effect of Microloans on Business owners' Revenues: Impact from Financial Constraint

This table presents the impact of microloans on migrant business owners' revenues compared to their native counterparts, conditional on business owners' financial constraint. We report the result of the regression in Equation (2). $FinConstraint_i$ is an indicator variable equal to one if a business owner face more financial constraint, and zero otherwise. In Column (1), we define "more financial constraint" business owners as those with initial revenue, $Revenue_{2017M3}$, below sample median. In Column (2), "more financial constraint" business owners are those with initial revenue, $Revenue_{2017M3}$, below 25th percentile. In Column (3), "more financial constraint" is defined in the same way as in Column (2) but the regression sample only includes business owner with initial revenue below 25th percentile or above 75th percentile. In Column (4), "more financial constraint" business owners are those with initial revenue, $Revenue_{2017M3}$, of the first tercile (below 33th percentile) and the sample only includes individuals with initial revenue of the first and third tercile (above 66th percentile). In Column (5), we define "more financial constraint" business owners as those only have a QR money-receiving code but do not own an online shop. These business owners are typically viewed as less sophisticated and having less resources. $Migrant_i$ is an indicator variable equal to one if business owner i is a migrant (his/her residential province as of the end of Dec. 2018 differs from the province where he/she was born), and zero otherwise. $DrawLoan_{it}$ is an indicator variable equal to one if business owner i draws loan credit during month t. Other variable definitions are in Online Appendix Table A1. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Dep. var. = *Ln(Revenue)*	Low Initial Revenue				(5)
	(1) <p50	(2) <p25	(3) Q1(vs Q4)	(4) Q1 (vs Q3)	QR code only
$FinConstraint_i$	-1.4064***	-1.4457***	-2.2000***	-1.9224***	-0.3238***
	[0.0150]	[0.0181]	[0.0224]	[0.0191]	[0.0352]
$Migrant_i$	0.0337**	-0.0102	-0.0017	0.0075	0.0794***
	[0.0147]	[0.0126]	[0.0206]	[0.0175]	[0.0135]
$DrawLoan_{it}$	0.1748***	0.1078***	0.1005***	0.2040***	0.1622***
	[0.0305]	[0.0266]	[0.0397]	[0.0342]	[0.0273]
$Migrant_i \times DrawLoan_{it}$	-0.1733**	0.0396	-0.0571	-0.1881***	0.0423
	[0.0418]	[0.0356]	[0.0579]	[0.0477]	[0.0366]
$Migrant_i \times FinConstraint_i$	-0.0289	0.0305	0.0309	0.0129	-0.3708***
	[0.0210]	[0.0250]	[0.0299]	[0.0253]	[0.0335]
$DrawLoan_{it} \times FinConstraint_i$	-0.0315	0.0871	0.1481**	0.0054	-0.3211***
	[0.0434]	[0.0530]	[0.0600]	[0.0501]	[0.0724]
$Migrant_i \times DrawLoan_{it}$	0.4672***	0.2252***	0.2519***	0.4472***	0.5389***
$\times FinConstraint_i$	[0.0602]	[0.0736]	[0.0864]	[0.0713]	[0.0962]
Controls	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
# Obs.	50,840	50,840	25,420	33,914	50,840
Adj R2	0.2717	0.2165	0.4604	0.4071	0.0348

Table 6: Effect of Microloans on Business owners' Revenues: Subsamples by Economic Development Level

This table conducts the same regressions as in Table 4 Column (1), subsampling by city economic development level. We use three ways to define economic development levels. The first is using city GDP per capita. All cities in matched sample are sorted based on their GDP per capita in year 2017 and then find terciles. The result on the cities with high vs. medium vs. low GDP per capita is reported in Columns (1), (2) and (3). The second is using Qinling Mountains-Huaihe River to divide China into the North and South regions. South regions are typically more economic developed. More specifically, we use the map of municipal central heating supply (which is based on Qinling Mountains-Huaihe River boundary). The cities with municipal heating supply coverage over 60% are classified as northern cities, and below 60% are southern cities. The result on the south versus the north is reported in Columns (4) and (5). In the second classification, we follow National Bureau of Statistics of China to define East area (more developed) and combine the rest of the areas as non-East area (less developed). The results based on East vs. non-East areas are in Columns (6)-(7). $Migrant_i$ is an indicator variable equal to one if business owner i is a migrant (his/her residential province as of the end of Dec. 2018 differs from the province where he/she was born), and zero otherwise. $DrawLoan_{it}$ is an indicator variable equal to one if business owner i draws loan credit during month t. Other variable definitions are in Online Appendix Table A1. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

	city GDP per capita			South vs. North		East vs. Non-East	
Dep. var. = *Ln(Revenue)*	(1) High	(2) Medium	(3) Low	(4) South	(5) North	(6) East	(7) non-East
$Migrant_i$	-0.0343**	0.1207**	0.0615**	-0.0161	0.1629***	0.0009	0.1145**
	[0.0153]	[0.0346]	[0.0251]	[0.0129]	[0.0380]	[0.0125]	[0.0559]
$DrawLoan_{it}$	0.0319	0.2758***	0.3692***	0.0552**	0.5769***	0.1048***	0.0874
	[0.0312]	[0.0648]	[0.0524]	[0.0268]	[0.0741]	[0.0260]	[0.0975]
$Migrant_i \times$	0.2074***	0.0551	-0.1078	0.1770***	-0.2086**	0.1362***	0.1352
$DrawLoan_{it}$	[0.0420]	[0.0926]	[0.0683]	[0.0361]	[0.0942]	[0.0348]	[0.1285]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# Obs.	32,054	6,138	12,648	46,314	4,526	48,918	1,922
Adj R2	0.0347	0.1211	0.0401	0.0268	0.1872	0.0302	0.2337

Table 7: Effect of Microloans on Business owners' Revenues: Subsamples by Risk Attitude

This table conducts the same regressions as in Table 4 Column (1), subsampling by risk attitude. We use two ways to define high and low risk attitude. The first is gender. Male represents high risk attitude and female represents low risk attitude. The second is business owner age. Younger (individuals whose age is lower than sample median) represents high risk attitude, and elder (individuals whose age is equal or higher than sample median) represents low risk attitude. $Migrant_i$ is an indicator variable equal to one if business owner i is a migrant (his/her residential province as of the end of Dec. 2018 differs from the province where he/she was born), and zero otherwise. $DrawLoan_{it}$ is an indicator variable equal to one if business owner i draws loan credit during month t. Other variable definitions are in Online Appendix Table A1. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Dep. var. = *Ln(Revenue)*	(1) Male	(2) Female	(3) Younger	(4) Elder
$Migrant_i$	0.0301*	-0.0605**	0.1511***	-0.1099**
	[0.0166]	[0.0188]	[0.0182]	[0.0168]
$DrawLoan_{it}$	0.1704***	0.0432	0.1273***	0.0939**
	[0.0324]	[0.0386]	[0.0351]	[0.0364]
$Migrant_i \times DrawLoan_{it}$	0.0623	0.1700***	-0.0153	0.2647***
	[0.0430]	[0.0543]	[0.0470]	[0.0490]
Controls	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
# Obs.	29,636	21,204	24,087	26,753
Adj R2	0.0394	0.0350	0.0546	0.0390

Table 8: Effect of Microloans on Business owners' Revenues: Subsamples by Signup Period

This table conducts the same regressions as in Table 4 Column (1), subsampling by signup period, the number of months since the business owner signs up with Alipay or Alipay's partners. "Sign up late" includes individuals whose signup_period at 2017M3 is smaller than sample median, and "Sign up early" group includes individuals whose signup_period is greater than or equal sample median. $Migrant_i$ is an indicator variable equal to one if business owner i is a migrant (his/her residential province as of the end of Dec. 2018 differs from the province where he/she was born), and zero otherwise. $DrawLoan_{it}$ is an indicator variable equal to one if business owner i draws loan credit during month t. Other variable definitions are in Online Appendix Table A1. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Dep. var. = *Ln(Revenue)*	Sign up early	Sign up late
$Migrant_i$	0.0551***	-0.0412*
	[0.0170]	[0.0179]
$DrawLoan_{it}$	0.1565***	0.0559
	[0.0345]	[0.0352]
$Migrant_i\times$	0.0541	0.1965***
$DrawLoan_{it}$	[0.0471]	[0.0468]
Controls	Yes	Yes
City FE	Yes	Yes
Industry FE	Yes	Yes
Occupation FE	Yes	Yes
Month FE	Yes	Yes
# Obs.	27,063	23,777
Adj R2	0.0443	0.0282

Internet Appendix

August 2021

Abstract

This online appendix presents additional tables analyzed in the paper. Table A1 includes the detailed definitions of all the variables used in the paper. Table A2 shows the correlation table among main variables in the matched sample used for regressions. Table A3 presents an overall positive relationship between (the city-level) mobile payment subdivision index and migrant population concentration.

Online Appendix Table A1
Variable Definitions

	Variable Type	Definition
Ln(Revenue)	Main dependent variable	Natural logarithm of the total amount of money from all recorded transactions that generate inflow into the Alipay account for each business owner each month. Revenues are in Chinese Yuan and winsorized at the 99% percentile.
Ln(Ntransactions)	Alternative dependent variable	Natural logarithm of the number of recorded transactions that generate inflow into the Alipay account for each business owner each month. *Ntransactions* is winsorized at the 99% percentile.
Ln(Ncustomers)	Alternative dependent variable	Natural logarithm of the number of unique customers recorded in transactions via Alipay account for each business owner each month (multiple payments received from one single payer is considered as one customer). *Ncustomers* is winsorized at the 99% percentile.
DrawLoan	Main independent variable	An indicator variable equal to one if the business owner draws loan credit during the month, and zero otherwise.
Migrant	Main independent variable	An indicator variable equal to one if the business owner is a migrant, and zero otherwise. An individual is classified as a migrant if his/her province of residence (as of the end of Dec. 2018) differs from the province in which he/she was born (i.e., the province shown on the individual's ID card). This variable is individual business owner invariant.
Taobao_seller	Control variable & Heterogeneity analysis	An indicator variable equal to one if an individual is an online business owner (i.e., a seller on the Taobao platform) as of the end of Dec. 2018, and zero otherwise.
IsQRcode	Control variable & Heterogeneity analysis	An indicator variable equal to one if an individual is a QR-code business owner (i.e., a business owner who has registered for and obtained a QR code for receiving business payments) as of the end of Dec. 2018, and zero otherwise.
City_degree	Control variable	A city-invariant variable on the city development level, with one being the most developed (e.g., Beijing, Shanghai) and six being the least developed.
Ln(gdppc)	Control variable & Heterogeneity analysis	Natural logarithm of GDP per capita for each city each year. The data is collected from city-level statistical yearbooks or statistical bulletins.
Farm_loan	Control variable	An indicator variable equal to one if, as of the end of Dec. 2018, an individual business owner has a separate microloan which is provided by Ant Group and specially used for farm-related business for people from rural areas.
Gender	Control variable	An individual business owner's gender, with 0 being female and 1 being male.
Age	Control variable	An individual business owner's age as of the end of Dec. 2018.

Age_sqr	Control variable	The square of the business owner's age
Edu_notmissing	Control variable	An indicator variable equal to one if the business owner's education information is not missing, and zero if the education information is missing. This variable is used instead of the original education variable because 90% of the sample does not have explicit education information.
Child_score	Control variable	The probability that the business owner has at least a child as of the end of Dec. 2018. The probability is estimated by Ant Group using big data analysis. This variable is individual business owner invariant.
RealEstate	Control variable	The probability that the business owner owns real property as of the end of Dec. 2018. The probability is estimated by Ant Group using big data analysis. This variable is individual business owner invariant.
Marriage	Control variable	The probability that the business owner is married as of the end of Dec. 2018. The probability is estimated by Ant Group using big data analysis. This variable is individual business owner invariant.
Signup_period	Control variable	The number of months since the business owner signs up with Alipay or Alipay's partners.
Occupation	Control variable	Occupation dummy variables on an individual's occupation.
Industry	Control variable	Industry dummy variables on the industry that the business is in.
FinConstraint	Heterogeneity Analysis	An indicator variable equal to one if a business owner face more financial constraint, and zero otherwise. We use several ways to define "more financial constraint": 1) initial revenue, $Revenue_{2017M3}$, below median of the entire matched sample; 2) initial revenue, $Revenue_{2017M3}$, below 25th percentile of the entire matched sample; 3) initial revenue, $Revenue_{2017M3}$, falling in the first tercile of the entire matched sample; 4) business owners who only have a QR money-receiving code and do not own any online shops.
CMPI	Appendix	City-level mobile payment sub-indexes in each month. These sub-indexes correspond to the usage of different mobile payment applications: 1) CMPI - offline goods purchase 2) CMPI - payments for public transportation 3) CMPI - payments for education 4) CMPI - payments for health care 5) CMPI - payments for utility bills 6) CMPI - payments to government agencies 7) CMPI - information inquiry on government-related matters We further calculate the average of the above seven sub-indexes to obtain an overall index of mobile payment development in each city. The CMPI indexes are available from Jan. 2017 to Dec. 2018.

Online Appendix Table A2
Correlation Matrix of Main Variables Used in Regressions

	Revenue	*Ntransactions*	*Ncustomers*	*Taobao _seller*	*IsQRcode*	*Farm_loan*	*Edu _notmissing*	*Gender*	*Age*	*Signup _period*	*Child _score*	*RealEstate*
$Ntransactions_{t-1}$	0.5421											
$Ncustomers_{t-1}$	0.5352	0.9650										
Taobao_seller	0.0373	-0.0471	0.0314									
IsQRcode	-0.0400	-0.0512	-0.0842	-0.2344								
Farm_loan	-0.034	0.0354	0.0314	-0.0205	-0.0394							
Edu_notmissing	-0.0121	-0.0078	-0.0115	-0.0059	-0.0088	0.0024						
Gender	-0.0237	0.0104	-0.0033	-0.1122	0.0870	0.0007	-0.0334					
Age	0.0383	0.0607	0.0644	-0.0494	-0.064	-0.0052	-0.0514	-0.0506				
Signup_period	0.0889	-0.0262	-0.0047	0.3069	-0.1368	-0.0267	-0.0240	-0.0370	0.0654			
Child_score	0.0200	0.0182	0.0207	-0.0458	-0.1316	0.0125	-0.0279	-0.0740	0.5197	0.0516		
RealEstate	-0.0168	0.0166	-0.0018	-0.1070	0.1491	-0.0252	0.0147	0.0876	0.0234	0.0645	-0.0772	
Marriage	-0.1074	-0.1529	-0.1721	-0.0970	0.1665	-0.0582	-0.0058	0.0553	-0.5396	0.0374	-0.3957	0.1825

Online Appendix Table A3
Migrant Population and Mobile Payment Development

This table presents the correlation between mobile payment index (*CMPI*) and city-level migrant population concentration. The table reports only the estimated coefficient of b_1 for the panel regression below

$$CMPI_{city,t} = b_0 + b_1 X + b_2 \ln\left(gdp_{city,yr}\right) + \varepsilon_{city,t}$$

$$where\ X = x_{city,\ yr} = \frac{migrant\ population_{city,yr}}{resident\ population_{city,yr}} \cdot 100$$

$$or\ \ X = \bar{x}_{city} = \frac{\overline{migrant\ population}_{city}}{\overline{resident\ population}_{city}} \cdot 100$$

The city-year level resident population, registered population, and GDP data are obtained from the local statistical yearbooks or statistical bulletins. Migrant population is measured by the difference between city-level resident population and the population of the people who were born in this city. Each of Columns (1)-(7) uses one CMPI sub-index. Specifically, (1) CMPI - offline goods purchase; (2) CMPI - payments for public transportation; (3) CMPI - payments for education; (4) CMPI - payments for health care; (5) CMPI - payments for utility bills; (6) CMPI - payments to government agencies; (7) CMPI - information inquiry on government-related matters. Panel A reports the coefficient of $x_{city,yr}$ (first definition of X above), and Panel B reports the coefficient of $\bar{x}_{city}$(second definition above, averaging city-year-level migrant and resident population separately over sample years of 2017-2018). Ln(gdp) is natural logarithm of GDP for each city each year. Robust standard errors are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels.

Panel A: City-level proportion of migrant population for each year as main independent variable

	Dep. Var.: 7 CMPI sub-indexes						
	(1) Offline purchase	(2) Public transport.	(3) Education	(4) Health care	(5) Utility bills	(6) Govt. payment	(7) Govt. inquiry
City-level % migrant population	0.2662***	0.3271***	0.4715***	0.3317***	-0.1609***	-0.0223	0.2932***
	[0.0129]	[0.0187]	[0.0384]	[0.0212]	[0.0229]	[0.0151]	[0.0247]
Control for Ln(gdp)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs.	5597	5497	5592	5592	5291	5592	5625
AdjR2	0.1246	0.2388	0.1715	0.2020	0.0065	0.0039	0.1528

Panel B: City-level proportion of migrant population (averaged over 2017-2018) as main independent variable

	Dep. Var.: 7 CMPI sub-indexes						
	(1) Offline purchase	(2) Public transport.	(3) Education	(4) Health care	(5) Utility bills	(6) Govt. payment	(7) Govt. inquiry
City-level % migrant population (avg. over sample years)	0.2741***	0.3311***	0.4954***	0.3487***	-0.1999***	-0.0312**	0.2976***
	[0.0123]	[0.0181]	[0.0372]	[0.0213]	[0.0237]	[0.0145]	[0.0247]
Control for Ln(gdp)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
#Obs.	5777	5677	5772	5772	5471	5772	5625
Adj R2	0.1271	0.2440	0.1689	0.2028	0.0089	0.0055	0.1507