This readme file describes the structure of the directories of the replication package containing the data (or relevant sources when using proprietary datasets) and programs needed to replicate the results in "The Macroeconomics of Microfinance." When necessary, readme files within the subdirectories provide additional information about the individual files and procedures used (we reproduce the information in these additional readme files below).

Data souces. This subdirectory contains the data and programs used to do the calculations in tables 1-4. We use Stata/SE 15.0 for Mac (64-bit Intel), Revision 06 Jun 2017. We run Stata in MacBook Pro (13-inch, 2017, Two Thunderbolt 3 ports) 2.5 GHz Intel Core i7.

- Table 1. This subdirectory contains the data and code used to produced Table 1.
    - Table1Numbers.do - This file produces the numbers in Table 1. It uses as an input the data file CombinedData2014.dta.
    - CombinedData2014.dta - The data file contains the following values for 2014: number of active borrowers (numberofactiveborrowers) and gross loan portfoliois (grossloan) from aggregating proprietary MiX data, this data was downloaded from the website of the Microfinance Information Exchange on February 20, 2017; population (pop), nominal GDP data (ngdpna), and real GDP per capita (rgdpcap) from PWT 9.0; and private credit from banks and other financial instittuionsas (pcrdbofgdp), private bond market capitalization (prbond), and public bond market capitalization (pubond), all as a % of GDP, from the Financial Structures Database, June 2016 (Beck et al., 2000). While we use an earlier proprietary version, the most updated MiX data is now publicly available online through an agreement with the World Bank, see World Bank Data Catalog, https://datacatalog.worldbank.org/dataset/mix-market..
    - Table1numbers.log - log file with the results obtained by running Table1Numbers.do.
- Table 2. This subdirectory contains the data and codes used to produced the target moments reported in Table 2. We reproduce below the information in the readme.txt file:
    - Table 2 - Top 10-percent income share -- This is the average between 2008 and 2015 of the numbers from the World Inequality Database, available at https://wid.world/country/india/. We choose the most recent and relative stable part of the sample.
    - Table 2 - Average Establishment Size.xlsx -- calculates average establishment size as the ratio of the total number of workers to the total number of establishments in India. The underlying data is from Tables A.1

and A.2 of the Appendix of the All India Report of Fifth Economic Census of India, Ministry of Statistics & Programme Implementation, 2008. The file Tables_All_India_Report_Of_Fifth_Economic_Census.pdf, which was downloaded from http://www.mospi.gov.in/all-india-report-fifth-economic-census, contains the original tables from the appendix of this report.

- Table 2 - External Finance to GDP.xls -- The raw data are the Financial Structures Database, 2010 as described in the 'About' worksheet. We calculate external finance to GDP as the sum of private credit from money banks and other financial institutions, the public bond market, and 35 percent of the public stock market (since we assume the remainder is the fraction that goes to capital gains rather than financing the firm per se).

- Table 2 - Interest rate in India.xlsx - This calculates the real interest rate as the difference between the savings rate and the inflation rate. Savings rates are the most relevant interest rate, but are first available in the 2000s. We therefore use the average of the years 2000-05 instead of the 1990s.

- Table2IncomeAutoCorrelation.do - This calculate the income autocorrelation using the ICRISAT.dta dataset, which is ICRISAT.dta and outputs the result in the file Table2IndiaIncomeAutoCorrelation.log.

- The concentration of firms and exit rate are not from disclosable sources:

    - Top 10-percentile employment share -- This moment is calculated using micro data from the 2005 Economic Census of India. In particular, we measure employment are hired workers and calculate the share of total hired workers accounted by the top 10% of firms ranked in terms of their number of hired workers.

    - Establishment exit rate -- 4.72% implied annual rate from 1994-1995 to 2010-2011 in the ASI-NSS combined as calculated using the dataset in Hsieh and Klenow, "The Life Cycle of Plants in India and Mexico," Quarterly Journal of Economics, 129(3): pp 1035-84. The number corresponds to the weighted average of the exit rate by age reported in Figure II and was reported to us by Pete Klenow (personal communication, April 11, 2015).

- Tables 3 and 4.

    - Data Moments for Tables 3 and 4.xlsx -- This file contains two worksheets where the data moments in Tables 3 and 4 are calculated, including the reported ratios, percentage growth, and confidence intervals. Cell comments explain the sources and calculations. The bolded numbers are those reported in the tables of our text.

- Table 3 Thai Data Moments: This worksheet contains the values taken from Kaboski and Townsend (AEJ Applied, 2012). For investment and investment probability, we take them from Kaboski and Townsend (2011), since aggregate investment numbers are not available in Kaboski and Townsend (2011). Since we want to rely only on published data, and treat things symmetrically with India, we assume no information on the variance-covariance terms, and so we take the denominator as given.

- Table 4 Indian Data Moments: This worksheet contains the values taken from Banerjee, Duflo, Glennerster, and Kinnan (AEJ Applied, 2015). Since we lack covariance terms in the variance-covariance matrix, we assume that the denominator is known.

Data souces. This subdirectory contains the codes used to run the simulations reported in the paper. Here we reproduce the information in the readme.txt file that can be found within this subdirectory:

> There are ten (10) files, including two (2) input files. We use Intel compiler and run the codes on 40 processors using Intel
> MPI.
>
> - "distmf40e280.dat": The probabilities of 40 discrete entrepreneurial productivity (z). Input file.
>
> - "supportmf40e280.dat": The values of 40 discrete entrepreneurial productivity (z). Input file.
>
> - "steadynomf40by2.f90": Computes the stationary equilibrium without microfinance.
>
> - "steadymf40by2_clspread.f90": Computes the stationary equilibrium with microfinance.
>
> - "lrsoe_clspread.f90": Computes the stationary equilibrium with microfinance for a small open economy taking the world interest rate as given.
>
> - "tranmf40by2_clspread.f90": Computes the transition dynamics following the introduction of microfinance. It takes as input "p.nomf", which is the initial wealth distribution (density) conditional on entrepreneurial and labor productivity (40 by 2), which can be constructed from the output file of "steadynomf40by2.f90".
>
> - "srpe_clspread_tha.f90": Generates the decision rule for the microfinance intervention in partial equilibrium for Thailand (Table 3). See

"mmt_srpe_clspr.m" below.

- "srpe_clspread_ind.f90": Generates the decision rule for the microfinance intervention in partial equilibrium for India (Table 4). See "mmt_srpe_clspr.m" below.

- "mmt_srpe_clspr.m": A MATLAB file using the output files from "steadynomf40by2.f90" and "srpe_clspread_tha/ind.f90", compute the changes in moments of interest in partial equilibrium reported in Tables 3 and 4.

- "welfarecomp.m": A MATLAB file using output files from "steadynomf40by2.f90" and "tranmf40by2_clspread.f90", construct Figure A.1.

- "twofer.f90": This is the file that generated the twosector results in our 2012 NBER working paper. A summary of the results can be found in Appendix C of the current paper.