Internet Appendix for "Dynamic Bank Expansion: Spatial Growth, Financial Access, and Inequality"*

Yan Ji Songyuan Teng Robert M. Townsend

Contents

*Internet Appendix for Ji, Teng and Townsend (2022). Ji: Department of Finance, Hong Kong University of Science and Technology (email: jiy@ust.hk); Teng: Department of Economics, Yale University (email: songyuan.teng@yale.edu); Townsend: Department of Economics, MIT (email: rtownsen@mit.edu).

1 Supplemental Information for Data

In this section of the appendix, we provide supplemental information on the data. Subsection 1.1 describes the road network and branch locations. Subsection 1.2 provides details on how we construct the spatial economic variables (i.e., values that vary across geographic locations) used in our empirical and quantitative analyses. Subsection 1.3 describes the construction of aggregate variables used for model calibration and quantitative analyses.

1.1 Road Network and Branch Locations

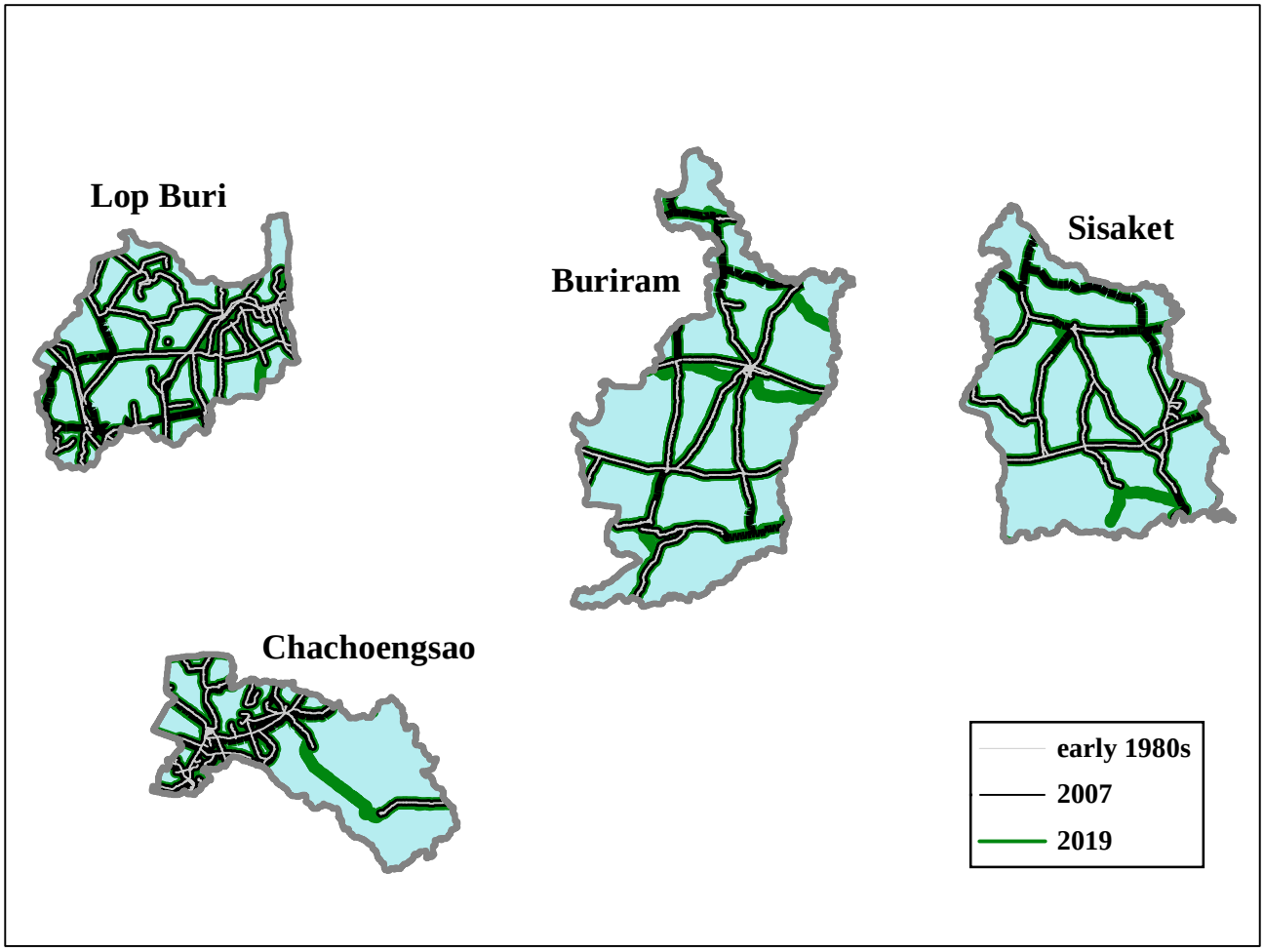
Road Network. The road network is constructed as in Felkner and Townsend (2011). We obtain highly detailed road network data from Royal Thai Government maps from the early 1980s to certify that all road networks were in place before the 1986 base year in our study. We estimate the average car travel speeds on different types of roads in Thailand using geographic information system (GIS) data (see Table OA.1). The speed data are used to estimate the car travel time along the road network between each pair of markets and between each village and branch location. We assign villages that are not located on any road to the nearest road and calculate the travel time based on a footpath travel speed of 15 km/hour.

Table OA.1: Road types and estimated average car travel speeds.

Code	Road type	Average speed (km/hour)
1	All weather: hard surface, two or more lanes wide	45
2	All weather: loose or light surface, two or more lanes wide	38
3	All weather: hard surface, one lane wide	38
4	All weather: loose or light surface, one lane wide	30
5	Fair or dry weather: loose surface	25
6	Cart track	20
7	Footpath, trail	15

The road network is fixed in our analyses. The fact that the road network is endogenous may raise potential concerns for our empirical and quantitative results. Specifically, economic development can lead to the construction of new roads, which in turn promotes credit access and economic growth in subsequent years. Felkner and Townsend (2011) compare the Royal Thai Government road data with more recent maps obtained from American Digital Cartography (ADC) as well as with current Thai road maps and Google Maps data.[1] They find that no new primary roads (highways and high-quality paved roads) are constructed after 1986: all such roads are present in the early 1980s in the Royal Thai Government map data. Some new secondary

[1]The ADC is rated as the best quality database available on global road networks by a World Bank study (Nelson, de Sherbinin and Pozzi, 2006), including in comparison with declassified United States Central Intelligence Agency data.

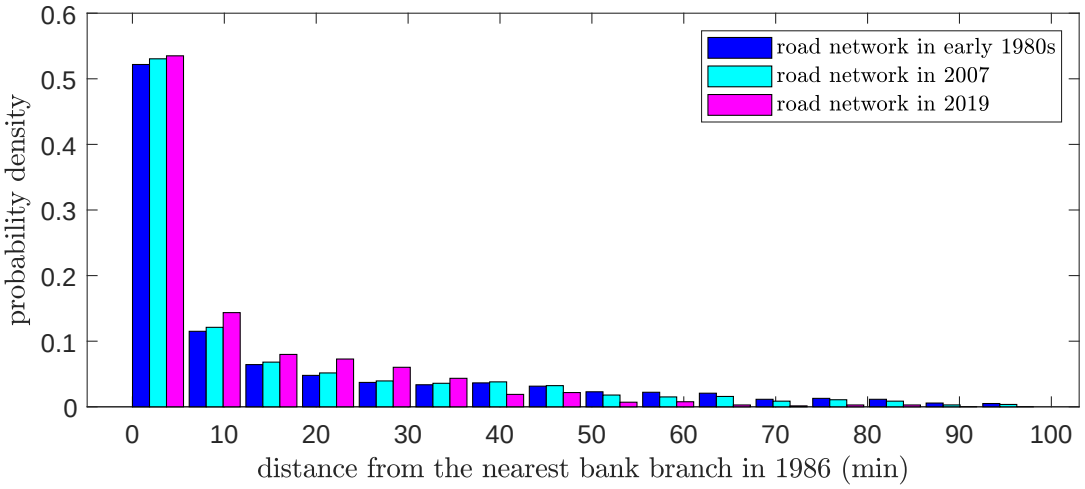


Note: The thinnest links represent the roads in the early 1980s, which are also used in our empirical and quantitative analyses; the links of intermediate width represent the roads in 2007, obtained from ADC; and the thickest links represent the roads in 2019, obtained from OpenStreetMap (https://download.geofabrik.de/asia/thailand.html).

Figure OA.1: Snapshots of road networks in four representative provinces in Thailand, 1980-2019.

roads were constructed, which are present in the ADC maps but not in the earlier Royal Thai Government maps. However, these new secondary roads are very small in number. Almost all were either relatively minor additions to market centers that already possessed extensive secondary road networks or were constructed in rural areas that did not experience enterprise growth. Felkner and Townsend (2011) regress local enterprise growth rates on the distance from new secondary roads and obtain a coefficient that is not statistically significant.

To provide some examples, Figure OA.1 presents snapshots of road networks (including primary and secondary roads only) in four representative provinces in Thailand. These same provinces are used in the Townsend Thai surveys. Lop Buri and Chachoengsao are in the more



Note: This figure plots the histograms of markets' distance from the nearest bank branch in 1986 based on car travel time using the road networks (including all types of roads) observed in the early 1980s, 2007, and 2019.

Figure OA.2: Distribution of markets' distances from the nearest bank branch in 1986.

highly developed central region located near Bangkok; and Sisaket and Buriram are in the poorer semi-arid northeast region. We present the road snapshots observed in different years. The thinnest links represent the roads in the early 1980s, which are those used in our empirical and quantitative analyses. The intermediate width links represent the roads in 2007, based on ADC data; and the thickest links represent the roads in 2019 based on OpenStreetMap (https://download.geofabrik.de/asia/thailand.html). The figure clearly shows that the road network in 2007 is quite similar to that in the early 1980s, with only a few minor additions. Comparing the road networks for 2007 and 2019, we can see that a few new roads were built during this period, but that most roads were already present in the early 1980s.

As an additional validation test, we calculate the car travel time for each market from the nearest bank branch in 1986 based on the road networks (including all types of roads) observed in the early 1980s, 2007, and 2019. Figure OA.2 shows that the three histograms are very close to each other. These car travel times between markets are the relevant metrics used in our model calibration and quantitative analyses. The average (median) car travel time from the nearest bank branch across all 1,428 markets is 14.8 (5.23) minutes based on the early-1980s road network, 14.0 (5.03) minutes based on the 2007 road network, and 13.4 (4.93) minutes based on the 2019 road network. Because our study period is from 1986 to 1996, which is about 10 years before 2007, it seems reasonable to assume that the roads in the early 1980s remain unchanged during our study period for the empirical and quantitative analyses.

Branch Location. We obtain the village/township (tambon)/municipality name of each branch location using a Google Maps application programming interface (API). For branches that can be

matched directly to a village, the village's location, which is represented as point data in the GIS, is used as a proxy for the branch's location. Branches that cannot be matched at the village level are matched to tambons and/or municipal districts (geo-units represented as polygonal areas in the GIS), and are assigned to road network intersections according to the following procedures:

(i) For each branch, find the region where the branch is located. If the branch is mapped to a tambon or a municipal district, the corresponding polygonal area is considered the associated region. If the branch is mapped to both a tambon and a municipal district, the overlap of the two polygons is considered the associated region.

(ii) For the entire economy, find all of the road intersections.

(iii) For each intersection, find all of the roads that are connected to it. We assign a weight to each type of road equal to 8 minus the road's code number (see Table OA.1). Thus, for example, the highest level of road with a code of 1 has the highest weight, 7. The weight of each intersection is calculated as the sum of all of the weights of the connecting roads. These weights represent the relative importance of intersections in terms of connecting major roads.

(iv) For each region defined in step (i), find all of the branches that belong to the region. The location of the first branch opened in this region is assigned to the intersection with the highest weight, the location of the second opened branch is assigned to the intersection with the second highest weight, and so on. We require branches opened by the same bank to be at least 500 meters apart.

1.2 Construction of Spatial Economic Variables

Primitive Market Heterogeneity. The three primitive market characteristics in Section 4 of the main text are constructed as follows. We estimate the population of each market by summing up the village and municipal populations in 1986. Village-level information is obtained from the Thai Community Development Department (CDD) data. CDD provides a complete village census (of over 60,000 villages) based on interviews with headmen, which has been collected by the Rural Development Committee biennially since 1986. Information about municipalities is collected from the Thai Population and Housing Census. The population of each village is mapped to its nearest market based on the road network. The population of each municipality is distributed equally among the markets within its administrative boundary. The population density within a market is equal to the market's population divided by the market's area computed using GIS data. Panel A of Figure OA.3 displays the estimated population density in 1986.

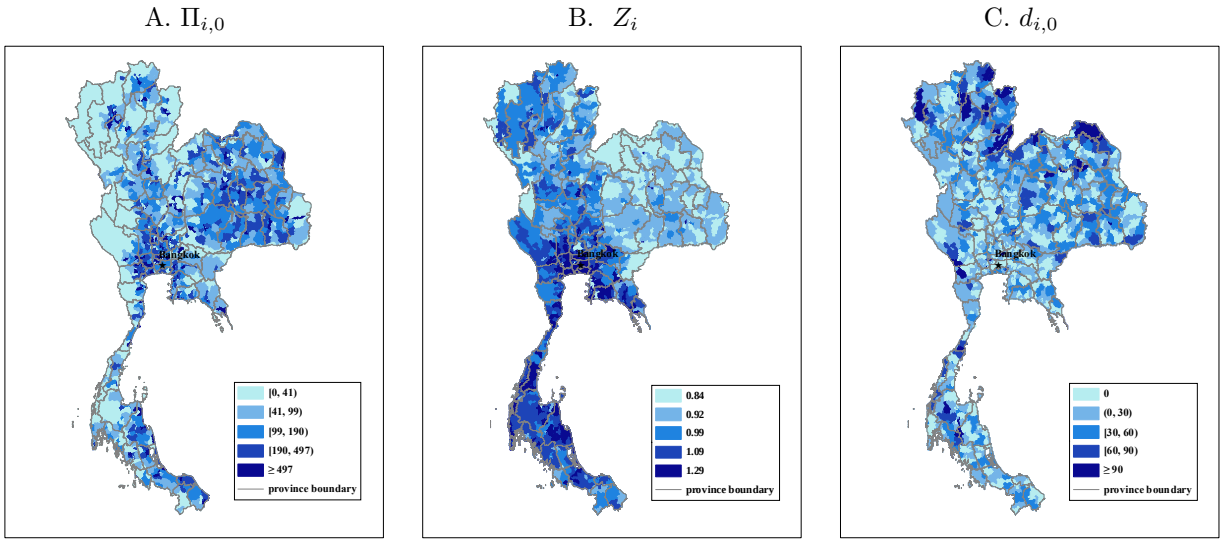
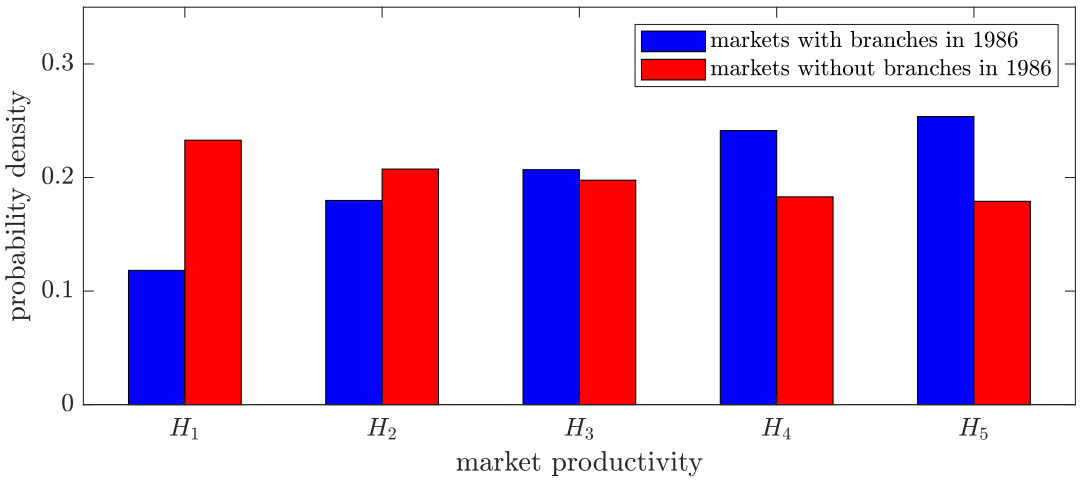


Figure OA.3: Primitive market characteristics in 1986.

The estimation of market-level productivity Z_i is explained in Section 4.1 of the main text. Panel B of Figure OA.3 displays the distribution of Z_i. Figure OA.4 shows that the markets with branches in 1986 tend to be more productive than other markets.

For each market, we estimate the distance from the nearest bank branch in 1986 based on the road network and branch location data (see Online Appendix 1.1). Panel C of Figure OA.3 displays the estimated distance from the nearest bank branch in 1986.

Nontargeted Outcome Variables. The outcome variables in our empirical analyses (Section 2.3 and Figure 5 of the main text) are constructed using the CDD data. For each village, CDD documents the number of households exclusively engaged in selling labor, which we regard as the number of households employed as wage earners. CDD also provides information on the number of households exclusively engaged in retail business and in the cottage industry. We follow Felkner and Townsend (2011) and define the number of households engaged in entrepreneurial activities as the sum of these two numbers. In addition, CDD documents the credit access conditions for each village, reporting on whether households have commercial bank loans. We obtain the income of each village from Townsend (2011). By aggregating all villages within each market, we construct market-level income per capita (defined as the market's total income divided by its population), employment (defined as the number of households employed as wage earners), the fraction of households with commercial bank loans (defined as the ratio of the number of households with commercial bank loans to the total number of households), and the fraction of entrepreneurs (defined as the ratio of the number of households engaged in entrepreneurial



Note: The blue bars represent the distribution of market productivity for markets with branches in 1986. The red bars represent the distribution of market productivity for markets without branches in 1986. The calibrated values are $\{H_1, H_2, ..., H_5\} = \{0.84, 0.92, 0.99, 1.09, 1.29\}$.

Figure OA.4: Productivity distribution of markets with and without branches in 1986.

activities to the total number of households). We obtain a biennial panel of 1,428 markets for the period from 1986 to 1996.

1.3 Construction of Aggregate Economic Variables

In Figure 8 of the main text, the time series of GDP growth are obtained from Jeong and Townsend (2007, 2008). The fraction of entrepreneurs with loans and the fractions of entrepreneurs, workers, and farmers are constructed using the Thai Socio-Economic Survey (SES), a nationally representative household-level survey covering over 11,000 households, which has been conducted biennially by the National Statistical Office since 1986. The credit access Gini is constructed using the market-level fraction of households with commercial bank loans from the CDD data, as described in Online Appendix 1.2. The computation of the Gini coefficient follows the standard formula. In Figure 9 of the main text, the aggregate income Gini coefficient, the top 10% share of total income, and the bottom 50% share of total income are constructed using the household-level income data provided by SES.

The migration rates are constructed using the CDD data. For each village, the CDD data document the number of people working outside the tambon; these people are considered out-migrants of the township. By aggregating the out-migrants and populations across all villages, we construct the out-migrant share in each survey year at the country level. In addition, based on the bank branch locations in each survey year, we construct the ratio of the share of out-migrants in markets with branches to the share of out-migrants in markets without branches. These moments

are used to calibrate parameters κ and η in Section 4 of the main text.

The data series presented in panel B of Figure 10 of the main text are constructed using market-level income, which is obtained by aggregating the income of each village as described in Online Appendix 1.2. The data series in panel A of Figure 10 of the main text are constructed using the variables "wages and salaries" and "profit, non-farm" from the SES data, which provide workers' wages and entrepreneurs' business incomes, respectively, at the village level. We first map the villages in the SES with those in the CDD data using the unique village ID. Then, we estimate workers' wages and entrepreneurs' business incomes for other villages included in the CDD data using a projection method. Finally, for each market, we calculate the income Gini coefficients using workers' wages and entrepreneurs' business incomes for all villages contained in the market.

The 1997 Townsend Thai annual survey (available at Harvard Dataverse) documents the amount of loans borrowed from commercial banks and the value of collateral for the loans at the household level. After filtering out unsecured loans (those that do not require collateral), we use these two variables to construct the loan-to-collateral ratio distribution. The median value of this distribution is used as an untargeted moment to verify the validity of our calibrated value of ξ in Section 4 of the main text.

We use the Townsend Thai monthly surveys from January 1999 to December 2011 to construct the autocorrelation of log revenue and the standard deviation of log revenue growth. These surveys have been conducted since August 1998 in 16 villages in Thailand (Townsend, Sakunthasathien and Jordan, 2013). The monthly surveys provide households' total revenue and revenue from labor. Following Samphantharak and Townsend (2018), we construct asset-generated revenue by taking the difference between total revenue and revenue from labor. Asset-generated revenue includes revenue from cultivation, livestock, fish and shrimp, business, and other, recorded as five separate variables in the monthly surveys. Then, we aggregate the monthly asset-generated revenue in each year to obtain yearly data for the 1999-2011 period. These variables are used to calibrate the parameters ρ and γ in Section 4 of the main text.

2 Supplemental Information for Empirical Analyses

In this section of the appendix, we provide supplemental information for our empirical analyses. Subsection 2.1 provides the implementation details for our empirical research design, including the difference-in-differences (DID) approach with matching (our baseline specification) and the synthetic control method. Subsection 2.2 presents and compares the empirical estimates from these methods to show the robustness of our results. Subsection 2.3 presents suggestive evidence for the relationship between bank branches and household cash holdings.

Table OA.2: DID with matching: Covariates of treatment and control groups

	(1)	(2)	(3)	(4)	(5)	(6)
	CEM		NNM		PSM (baseline)	
	treatment	control	treatment	control	treatment	control
Log population	7.87	7.71	8.00	7.64	7.99	7.72
	(1.23)	(1.17)	(1.17)	(1.10)	(1.15)	(1.10)
Fraction of households with pickup trucks	0.07	0.07	0.07	0.07	0.07	0.08
	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.08)
Type of cooking fuel	0.31	0.30	0.31	0.34	0.31	0.37
	(0.38)	(0.38)	(0.37)	(0.39)	(0.37)	(0.41)
Distance from the nearest bank branch	25.02	22.07	24.93	19.71	24.61	19.70
	(23.13)	(21.33)	(21.31)	(20.34)	(20.59)	(21.45)
No. of branch openings	362	362	457	457	462	462

Note: This table presents the summary statistics of matched treatment and control markets after applying CEM (columns (1) and (2)), NNM (columns (3) and (4)), and PSM (columns (5) and (6)). Standard deviations are reported in parentheses. The mean and standard deviation of each covariate is computed based on the values 2 years prior to the treatment, across all pairs of treatment and control markets.

2.1 Implementation Details

In Subsection 2.1.1, we describe three matching methods, including coarsened exact matching (CEM), nearest neighbor matching (NNM), and our baseline matching method, propensity score matching (PSM). These methods are used to construct matched pairs of treatment and control markets based on which we apply DID (i.e., equation (2.1) in the main text). In Subsection 2.1.2, we present the synthetic control method.

2.1.1 DID with Matching Methods

This section presents the implementation procedure for DID with matching methods. Although our baseline specification uses PSM, we use alternative matching methods (i.e., CEM and NNM) to check the robustness of our results. We present the covariates of treatment and control groups in Table OA.2. It is evident that the treatment and control groups have roughly balanced means and standard deviations for all covariates used in our DID with different matching methods.

CEM. CEM is a nonparametric matching method and thus only a small number of covariates can be chosen due to the "curse of dimensionality" (e.g., Iacus, King and Porro, 2012).[2] CEM ensures that the covariates of matched treatment and control markets are balanced, but some treatment markets may not be matched with a control market.

[2]For the applications of DID with matching based on CEM, see, e.g., Balsmeier, Fleming and Manso (2017) and Jager and Heining (2019).

Starting from year $t = 1986$, we implement CEM year by year according to the following steps:

(i) Define the candidate pool, which includes the treatment markets with a new branch opened in year $t-1$ or t and the control markets in the donor pool of year t, where the donor pool is defined in the main text.

(ii) Construct the joint distribution of the four covariates in year $t-2$, based on markets in the candidate pool.[3] Find cutoff values for the tertiles of the marginal distribution of each covariate. Based on these cutoff values, coarsen the joint distribution of the four covariates into $81 (= 3^4)$ strata.

(iii) Assign each market in the candidate pool to a stratum by matching the values of all four covariates of the market in year $t-2$ with those of the stratum.

(iv) Focus on the strata that contain both treatment and control markets. For each stratum, if there is only one treatment market, we follow Jager and Heining (2019) and select the control market that has a propensity score closest to that of the treatment market.[4] If there are multiple treatment markets, we first list all treatment markets in descending order according to the propensity score. Then, we sequentially match each treatment market in the list with a control market that has the closest propensity score. This matching process is conducted with no replacement; it stops if the remaining markets in the stratum are either all treatment or all control markets. This implies that some treatment markets are not matched in the end, a common issue with CEM.

(v) Construct the donor pool of year $t+2$, which is the donor pool of year t excluding all matched control markets in step (vi).

(vi) Move to year $t+2$ and repeat steps (i) to (v).

NNM. As noted above, CEM ensures that the covariates of treatment and matched control markets are balanced, but some treatment markets may not be matched with a control market. NNM matches each treatment market to its closest market from a donor pool based on the Mahalanobis distance (e.g., Rubin, 1980).[5] It guarantees that each treatment market is matched with a control market.

[3] We match each treatment market based on the value of covariates in the period prior to the treatment, following common practice in the literature (e.g., Jager and Heining, 2019; Smith et al., 2019; Bennedsen, Pérez-González and Wolfenzon, 2020).

[4] The calculation of the propensity score is described below for the implementation of PSM. Our data for covariates start from 1986, and thus, for $t = 1986$, we construct the joint distribution of covariates using the 1986 data.

[5] For the applications of DID with matching based on NNM, see, e.g., Heckman, Ichimura and Todd (1997), Atanasov, Ivanov and Litvak (2012), Campello and Larrain (2016), and Almeida et al. (2017).

Starting from year $t = 1986$, we implement NNM year by year according to the following steps:

(i) Define the candidate pool, which includes the treatment markets in which a new bank branch is opened in year $t-1$ or t and the control markets in the donor pool of year t.

(ii) List all treatment markets in the candidate pool in random order.

(iii) Within the candidate pool, match each treatment market with a control market that has the smallest Mahalanobis distance from the treatment market. The Mahalanobis distance between any two markets i and j is given by $\sqrt{(x_i - \mu)' S^{-1} (x_j - \mu)}$, where x_i and x_j represent the vectors of the four covariates of markets i and j, respectively, in year $t-2$; and μ and S represent the mean vector and covariance matrix, respectively, of the four covariates in year $t-2$, across all markets in the candidate pool. This matching process is performed sequentially based on the random-order list constructed in step (ii), with no replacement.

(iv) Construct the donor pool of year $t+2$, which is the donor pool of year t excluding all matched control markets in step (iii).

(v) Move to year $t+2$ and repeat steps (i) to (iv).

We repeat the above steps 10,000 times and select the matching result that has the smallest sum for the Mahalanobis distance across all matched pairs of treatment and control markets.

PSM (Our Baseline Specification). PSM implements matching based on each market's single propensity score (e.g., Rosenbaum and Rubin, 1983).[6] It complements NNM by discovering pairs of comparable markets that potentially differ in their covariates.

Starting from year $t = 1986$, we implement PSM year by year according to the following steps:

(i) Define the candidate pool, which includes the treatment markets with a new branch opened in year $t-1$ or t and the control markets in the donor pool of year t.

(ii) Estimate the propensity score for each market in the candidate pool using the treatment dummies and covariates in year $t-2$ based on a logistic regression. In particular, we estimate a logistic model as follows:

$$\Pr(Treated_{i,t} = 1 | X_{i,t-2}) = \frac{1}{1 + \exp(-\beta X_{i,t-2})}, \tag{2.1}$$

[6]For the applications of DID with matching based on PSM, see, e.g., Smith and Todd (2005) and Couch and Placzek (2010).

where the dummy variable $Treated_{i,t}$ equals 1 if market i is treated in year $t-1$ or t. $X_{i,t-2}$ is a vector that includes a constant and the four covariates in year $t-2$. The logistic model is estimated using maximum likelihood estimation as follows:

$$\widehat{\beta} = \arg\max_{\beta} \sum_{i=1}^{n_t} [Treated_{i,t} \ln(\Pr(Treated_{i,t} = 1|X_{i,t-2})) + (1 - Treated_{i,t}) \ln(1 - \Pr(Treated_{i,t} = 1|X_{i,t-2}))], \tag{2.2}$$

where n_t is the number of markets in the candidate pool defined in step (i). The propensity score for each market in the candidate pool is the fitted value based on the covariates in year $t-2$ before the treatment year (i.e., $t-1$ or t):

$$\text{Propensity score}_{i,t-2} = \frac{1}{1 + \exp(-\widehat{\beta} X_{i,t-2})}. \tag{2.3}$$

(iii) List all treatment markets in the candidate pool in descending order according to the estimated propensity score in step (ii).

(iv) Within the candidate pool, match each treatment market with a control market that has the closest propensity score. This matching process is performed sequentially based on the descending-order list above, with no replacement.[7]

(v) Construct the donor pool of year $t+2$, which is the donor pool of year t excluding all matched control markets in step (iv).

(vi) Move to year $t+2$ and repeat steps (i) to (v).

2.1.2 Synthetic Control Method

This section presents the implementation procedure for the synthetic control method and cross-validation. The idea of the synthetic method is that a combination of control units rather than a single control unit often provides a better match for a treatment unit.[8] Thus, the synthetic control method can further improve the balance between the treatment and control groups (e.g., Abadie and Gardeazabal, 2003; Abadie, Diamond and Hainmueller, 2010).

Specifically, for each year t, we construct a synthetic control market for each treatment market that has a bank branch opened in that year. The synthetic control market is a weighted average

[7] In general, the descending-order list yields better matches than a random-order list (Rubin, 2006). We drop a matched pair of treatment and control markets if the difference between their propensity scores is larger than 0.1.

[8] For example, Abadie and Gardeazabal (2003) assign weights to two Spanish regions to approximate the economic growth that the treatment region, the Basque Country, would have experienced in the absence of terrorism.

of the markets in the donor pool. The weights are calculated such that the difference in the value of the covariates and the lagged outcome variables between the treatment market and the synthetic control market prior to year t is minimized. After matching each treatment market with a synthetic control market, we obtain a treatment effect estimate for each pair of treatment and synthetic control markets. The average treatment effect is given by the simple average across the estimates of all pairs.

We present the summary statistics for the covariates of the treatment and synthetic control groups in Table OA.3. The means and standard deviations are balanced for all four outcome variables.

Implementation Procedure for the Synthetic Control Method. We follow the procedure proposed by Abadie, Diamond and Hainmueller (2010). Starting from year $t = 1988$, we implement the synthetic control method year by year according to the following steps:

(i) For each treatment market i with a new bank branch opened in year $t-1$ or t, construct a synthetic control market, which is a weighted average of the control markets in the donor pool. Specifically, we find a vector of weights W_i over all control markets in the donor pool to minimize the discrepancy

$$\min_{W_i} \sqrt{(X_i - X_0 W_i)' V (X_i - X_0 W_i)},$$

which captures the difference between the treatment and synthetic control markets. The vector X_i includes all predictor variables, which are the four covariates used for PSM and the outcome variables in year $t-2$. The matrix X_0 captures the values of these predictors in all control markets. Thus, the vector $X_0 W_i$ captures the predictor variable in the synthetic control market. The matrix V is a positive semidefinite matrix that governs the relative contribution of each predictor to the discrepancy. We choose V as the identity matrix. The results are robust if we choose V as the covariance matrix of the predictor variables, across all markets treated in years $t-1$ and t and those in the donor pool. Following Abadie, Diamond and Hainmueller (2010), we restrict the elements of the weight vector W_i to be nonnegative and to sum to 1 to avoid extrapolation.

(ii) For each pair of treatment and synthetic control markets, obtain an estimate of the treatment effects (leading terms) by taking the difference in the outcome variable of years t, $t+2$, and $t+4$ ($t-6$, $t-4$, and $t-2$).

(iii) Move to year $t+2$ and repeat steps (i) to (ii).

(iv) The estimates of the average treatment effects are given by the simple average of the

Table OA.3: Synthetic control: Covariates of treatment and synthetic control groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln(Employment)		Frac. of HH with loans		ln(Income per capita)		Frac. of entrepreneurs	
	treatment	control	treatment	control	treatment	control	treatment	control
Log population	8.11	8.07	8.08	8.04	8.10	8.06	8.19	8.16
	(1.12)	(1.06)	(1.15)	(1.11)	(1.13)	(1.08)	(1.07)	(1.04)
Fraction of households with pickup trucks	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.06
	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)
Type of cooking fuel	0.32	0.32	0.31	0.31	0.34	0.34	0.30	0.30
	(0.37)	(0.35)	(0.37)	(0.35)	(0.37)	(0.35)	(0.35)	(0.34)
Distance from the nearest bank branch	24.53	24.53	24.45	24.40	23.88	23.88	24.76	24.75
	(20.32)	(20.27)	(20.27)	(20.20)	(20.20)	(20.15)	(20.29)	(20.27)
No. of branch openings	394	394	397	397	341	341	351	351

Note: This table presents the summary statistics of the treatment and synthetic control markets after applying the synthetic control method to estimate the impact of branch openings on employment, income per capita, the fraction of entrepreneurs and the fraction of households with bank loans. Standard deviations are reported in parentheses. The mean and standard deviation of each covariate is computed based on the value of the covariate 2 years prior to treatment, across all pairs of treatment and synthetic control markets.

treatment effects estimated in step (ii), across all pairs of treatment and synthetic control markets.

Cross-Validation. The cross-validation procedure follows Abadie, Diamond and Hainmueller (2015).

(i) For each year $t = 1990, ..., 1996$ and each treatment market i with a new branch opened in year $t-1$ or t, choose year $t-4$ as the training period and $t-2$ as the validation period.[9]

(ii) Given any positive semidefinite and diagonal matrix V, we can solve for the optimal weight $W_i(V)$ of treatment market i over all control markets in the donor pool as in step (i) of the implementation procedure for the synthetic control method, except that the vector X_i now only includes the predictor variables in the training period, that is, the four covariates and the outcome variable in year $t-4$.

(iii) Using the weights $W_i(V)$, construct the outcome variable for the synthetic control market of market i in year $t-2$, the validation period.

(iv) Calculate the difference in the outcome variable between the treatment market i and its corresponding synthetic control market in the validation period.

(v) Calculate the mean squared prediction error, which is the average squared difference (obtained from step (iv)) across all pairs of treatment and synthetic control markets in the sample period.

(vi) Repeat steps (i) to (v) to find the matrix V^* that minimizes the mean squared prediction error. Intuitively, V^* minimizes the out-of-sample prediction error in the validation period.

(vii) Use V^* to construct the synthetic control market for each treatment market based on the predictor variables in the validation period. The remaining steps follow steps (ii) to (iv) of the synthetic control method.

2.2 Robustness of Empirical Results

Our baseline empirical research design is DID with PSM (see Table 1 in Section 2.3 of the main text). As a robustness check, in Tables OA.4 to OA.7, we present the estimates from alternative methods. Specifically, columns (1) to (3) of each table present the estimates from DID with PSM (our baseline specification), CEM, and NNM, respectively (see Section 2.1.1).

[9]The length of the training and validation periods needs to be the same, as noted by Abadie, Diamond and Hainmueller (2015). Because of the short time span of our sample, we only use one-year data for training and validation.

Column (4) of each table presents the estimate from the synthetic control method. The synthetic control method is susceptible to overfitting issues because the characteristics of the treatment markets could be artificially matched by combining idiosyncratic variations of the markets in the donor pool. As suggested by Abadie, Diamond and Hainmueller (2015), we conduct two robustness checks to alleviate overfitting concerns. In column (5) of each table, we restrict the size of the donor pool and only consider markets similar to the treatment market when constructing the weights of the synthetic control market.[10] In column (6) of each table, we perform cross-validation by dividing the pretreatment years into a training period and a validation period (see Section 2.1.2 for details). Both robustness checks produce results similar to the synthetic control estimates in column (4). Finally, we conduct a placebo test in column (7) of each table to assess the significance of our estimates. The idea is to assess whether the statistically significant estimates for the impact of branch openings could be entirely due to chance. Following Abadie, Diamond and Hainmueller (2010), we apply the synthetic control method to randomly selected markets from the donor pool, which consists of markets that do not have any branch openings between 1987 and 1996.[11] Compared with the estimates of α_0, α_1, and α_2 in column (4) of each table, our placebo test produces very small and statistically insignificant estimates. This further reinforces our baseline findings that branch openings have significant impacts on the four outcome variables of interest.

Overall, Tables OA.4 to OA.7 indicate that our DID with different matching methods and the synthetic control method produce results of similar economic magnitude. Although none of the methods perfectly addresses the identification issue, the similarity in the results bolsters our confidence that our main identification strategy based on DID with PSM does estimate the effect of opening bank branches on local economic variables.

2.3 Deposit-Cash Ratios and Distance from the Nearest Branch

CDD does not provide information on savings or cash holdings. Thus, we cannot study the impact of branch openings on cash holdings using the DID approach in Section 2.3 of the main text. In this section of the appendix, we provide suggestive evidence of this impact using data on households' savings and cash holdings from the Townsend Thai monthly surveys. Because households living in the same villages are at an identical distance from the nearest bank branch, we aggregate the ratios of household savings in the bank to cash at the village level. This provides

[10] Specifically, for each treatment market, we only use 40 markets from the donor pool to construct the weights. These 40 markets have the shortest Mahalanobis distances to the treatment market.

[11] Specifically, the number of randomly selected markets in our placebo test is the same as the number shown in column (4) in each table. For each set of randomly selected markets, we apply the synthetic control method to estimate the treatment effect. Column (7) of each table reports the estimated treatment effect averaged across 10,000 independent random selections.

Table OA.4: Impact of opening bank branches on log employment.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DID with matching			Synthetic control			
	PSM	CEM	NNM	baseline	restricted donor pool	cross validate	placebo test
α_{-3}	−0.043	−0.119	−0.095	0.094	0.064	0.029	0.079
	(0.103)	(0.120)	(0.099)	(0.061)	(0.059)	(0.063)	(0.061)
α_{-2}	−0.013	−0.044	−0.044	0.092	0.063	0.078	0.144
	(0.097)	(0.116)	(0.093)	(0.060)	(0.060)	(0.061)	(0.060)
α_{-1}	0	0	0	0.012	−0.010	0.015	0.001
				(0.007)	(0.021)	(0.009)	(0.007)
α_0	0.186	0.160	0.219	0.168	0.173	0.213	0.090
	(0.092)	(0.102)	(0.091)	(0.071)	(0.066)	(0.075)	(0.071)
α_1	0.305	0.310	0.291	0.313	0.288	0.342	0.066
	(0.104)	(0.121)	(0.107)	(0.085)	(0.081)	(0.085)	(0.085)
α_2	0.341	0.319	0.323	0.304	0.351	0.290	0.058
	(0.115)	(0.127)	(0.113)	(0.090)	(0.095)	(0.139)	(0.090)
No. of branch openings	423	335	427	394	397	332	−−
No. of observations	5,076	4,020	5,124	4,728	4,764	3,984	4,728

Note: This table presents estimates for the impact of branch openings on market-level log employment. Columns (1) to (3) present the results of the DID estimation with different matching methods, PSM (baseline), CEM, and NNM. All coefficients are normalized relative to $\tau = -1$, 2 years prior to the treatment. Columns (4) presents the results of the synthetic control method. Columns (5) to (7) perform robustness checks for the synthetic control method. In column (5), we only consider markets similar to the treatment market when constructing the weights of the synthetic control market. In column (6), we apply the cross-validation procedure proposed by Abadie, Diamond and Hainmueller (2015). In column (7), we conduct a placebo test by applying the synthetic control method to randomly selected markets from the donor pool, which consists of markets that do not have any branch openings between 1987 and 1996. For columns (1) to (3), standard errors reported in parentheses are clustered at the market level to address potential concerns of serial correlation of outcomes across periods (Bertrand, Duflo and Mullainathan, 2004). For columns (4) to (7), standard errors are calculated following Arkhangelsky et al. (2021, algorithm 4).

Table OA.5: Impact of opening bank branches on log income per capita.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DID with matching			Synthetic control			
	PSM	CEM	NNM	baseline	restricted donor pool	cross validate	placebo test
α_{-3}	−0.071	−0.082	−0.087	−0.118	−0.036	−0.096	−0.008
	(0.104)	(0.125)	(0.105)	(0.061)	(0.055)	(0.060)	(0.061)
α_{-2}	−0.009	−0.015	−0.025	−0.105	−0.108	−0.087	−0.041
	(0.088)	(0.100)	(0.089)	(0.059)	(0.058)	(0.063)	(0.059)
α_{-1}	0	0	0	0.001	0.020	0.006	0.004
				(0.007)	(0.019)	(0.010)	(0.007)
α_0	0.174	0.202	0.203	0.159	0.146	0.157	0.037
	(0.089)	(0.097)	(0.084)	(0.050)	(0.049)	(0.058)	(0.050)
α_1	0.294	0.330	0.313	0.271	0.354	0.352	−0.006
	(0.094)	(0.102)	(0.090)	(0.056)	(0.076)	(0.089)	(0.056)
α_2	0.324	0.358	0.307	0.290	0.273	0.312	0.101
	(0.102)	(0.110)	(0.101)	(0.066)	(0.069)	(0.098)	(0.066)
No. of branch openings	400	325	400	341	344	306	−−
No. of observations	4,000	3,250	4,000	3,410	3,440	3,060	3,410

Note: This table presents estimates for the impact of branch openings on market-level log income per capita. Columns (1) to (3) present the results of the DID estimation with different matching methods, PSM (baseline), CEM, and NNM. All coefficients are normalized relative to $\tau = -1$, two years prior to the treatment. Columns (4) presents the results of the synthetic control method. Columns (5) to (7) perform robustness checks for the synthetic control method. In column (5), we only consider markets similar to the treatment market when constructing the weights of the synthetic control market. In column (6), we apply the cross-validation procedure proposed by Abadie, Diamond and Hainmueller (2015). In column (7), we conduct a placebo test by applying the synthetic control method to randomly selected markets from the donor pool, which consists of markets that do not have any branch openings between 1987 and 1996. For columns (1) to (3), standard errors reported in parentheses are clustered at the market level to address potential concerns of serial correlation of outcomes across periods (Bertrand, Duflo and Mullainathan, 2004). For columns (4) to (7), standard errors are calculated following Arkhangelsky et al. (2021, algorithm 4).

Table OA.6: Impact of opening bank branches on the fraction of entrepreneurs.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DID with matching			Synthetic control			
	PSM	CEM	NNM	baseline	restricted donor pool	cross validate	placebo test
α_{-3}	−0.012	−0.010	−0.015	−0.009	−0.014	−0.008	0.009
	(0.016)	(0.020)	(0.015)	(0.016)	(0.013)	(0.012)	(0.016)
α_{-2}	−0.001	0.004	−0.005	0.012	−0.001	0.014	0.005
	(0.015)	(0.018)	(0.015)	(0.013)	(0.012)	(0.012)	(0.013)
α_{-1}	0	0	0	0.004	0.006	0.008	0.003
				(0.004)	(0.008)	(0.004)	(0.004)
α_0	0.023	0.027	0.026	0.028	0.028	0.025	0.005
	(0.013)	(0.015)	(0.013)	(0.013)	(0.012)	(0.011)	(0.013)
α_1	0.035	0.035	0.035	0.043	0.038	0.039	0.004
	(0.016)	(0.018)	(0.015)	(0.015)	(0.014)	(0.016)	(0.015)
α_2	0.040	0.036	0.044	0.044	0.045	0.040	0.010
	(0.015)	(0.017)	(0.015)	(0.014)	(0.012)	(0.018)	(0.014)
No. of branch openings	393	302	404	351	351	293	−−
No. of observations	4,716	3,624	4,848	4,212	4,212	3,516	4,212

Note: This table presents estimates for the impact of branch openings on market-level fraction of entrepreneurs. Columns (1) to (3) present the results of the DID estimation with different matching methods, PSM (baseline), CEM, and NNM. All coefficients are normalized relative to $\tau = -1$, 2 years prior to the treatment. Columns (4) presents the results of the synthetic control method. Columns (5) to (7) perform robustness checks for the synthetic control method. In column (5), we only consider markets similar to the treatment market when constructing the weights of the synthetic control market. In column (6), we apply the cross-validation procedure proposed by Abadie, Diamond and Hainmueller (2015). In column (7), we conduct a placebo test by applying the synthetic control method to randomly selected markets from the donor pool, which consists of markets that do not have any branch openings between 1987 and 1996. For columns (1) to (3), standard errors reported in parentheses are clustered at the market level to address potential concerns of serial correlation of outcomes across periods (Bertrand, Duflo and Mullainathan, 2004). For columns (4) to (7), standard errors are calculated following Arkhangelsky et al. (2021, algorithm 4).

Table OA.7: Impact of opening bank branches on the fraction of households with bank loans.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	DID with matching			Synthetic control			
	PSM	CEM	NNM	baseline	restricted donor pool	cross validate	placebo test
α_{-3}	−0.006	−0.014	−0.019	−0.002	−0.013	−0.014	0.013
	(0.012)	(0.015)	(0.012)	(0.010)	(0.010)	(0.009)	(0.010)
α_{-2}	−0.007	−0.005	−0.000	0.001	−0.006	−0.005	0.012
	(0.011)	(0.013)	(0.011)	(0.009)	(0.009)	(0.008)	(0.009)
α_{-1}	0	0	0	−0.004	−0.002	−0.001	0.002
				(0.003)	(0.006)	(0.002)	(0.003)
α_0	0.034	0.035	0.029	0.036	0.032	0.039	0.007
	(0.009)	(0.010)	(0.009)	(0.008)	(0.013)	(0.009)	(0.008)
α_1	0.039	0.044	0.037	0.037	0.031	0.039	0.009
	(0.011)	(0.012)	(0.011)	(0.010)	(0.012)	(0.013)	(0.010)
α_2	0.038	0.044	0.034	0.045	0.035	0.041	0.014
	(0.012)	(0.013)	(0.011)	(0.014)	(0.011)	(0.014)	(0.014)
No. of branch openings	434	342	434	397	401	336	−−
No. of observations	5,208	4,104	5,208	4,764	4,812	4,032	4,764

Note: This table presents estimates for the impact of branch openings on market-level fraction of households with bank loans. Columns (1) to (3) present the results of the DID estimation with different matching methods, PSM (baseline), CEM, and NNM. All coefficients are normalized relative to $\tau = -1$, 2 years prior to the treatment. Columns (4) presents the results of the synthetic control method. Columns (5) to (7) perform robustness checks for the synthetic control method. In column (5), we only consider markets similar to the treatment market when constructing the weights of the synthetic control market. In column (6), we apply the cross-validation procedure proposed by Abadie, Diamond and Hainmueller (2015). In column (7), we conduct a placebo test by applying the synthetic control method to randomly selected markets from the donor pool, which consists of markets that do not have any branch openings between 1987 and 1996. For columns (1) to (3), standard errors reported in parentheses are clustered at the market level to address potential concerns of serial correlation of outcomes across periods (Bertrand, Duflo and Mullainathan, 2004). For columns (4) to (7), standard errors are calculated following Arkhangelsky et al. (2021, algorithm 4).

Table OA.8: Relationships between the distance from the nearest branch and deposit-cash ratios.

	(1)	(2)	(3)	(4)
Distance from the nearest branch	−1.16	−0.99	−1.08	−1.07
	(0.06)	(0.15)	(0.12)	(0.13)
Income level controls			√	√
Industry mix controls			√	√
Lagged growth				√
Year fixed effect		√	√	√
Township fixed effect		√	√	√
Observations	2,576	2,576	2,576	2,000
R^2	0.096	0.198	0.416	0.447

Note: We use the Townsend Thai data for the period from August 1998 to December 2011. The dependent variable is the village-level deposit-cash ratios (in percentage unit), constructed by aggregating the ratios of households' savings in the bank to cash for each village. The key independent variable of interest is the village's distance from the nearest bank branch. Income level controls include the village's asset, wealth, liability, contributed capital, cumulated savings, and net income. Industry mix controls include the village's income from different sources including cultivation, labor, livestock, fish and shrimp, and business. Lagged growth includes the village's 1-year, 2-year, and 3-year lagged growth rates of assets and wealth. Robust standard errors are reported in parentheses.

us with a panel of village-level deposit-cash ratios for the period from August 1998 to December 2011.

We provide suggestive evidence of the impact of branch openings on cash holdings by regressing the village-level deposit-cash ratios on the villages' distance from the nearest bank branch. Column (1) of Table OA.8 shows that a one-minute increase in car travel time (i.e., about 0.7 kilometers) is associated with approximately a 1% decrease in the deposit-cash ratio. In column (2), we further control for township and year fixed effects to capture the variation in the distance from the nearest bank branch over time caused by new branch openings. The results remain statistically significant and have comparable economic magnitudes.

The estimates in column (2) can be subject to omitted variable bias. For example, moving further away from economic centers (where branches are located), the industry mix will change from manufacturing/services to agriculture and therefore income levels will also change. The industry mix and income levels are correlated with villages' distance from the nearest branch and can influence the deposit-cash ratio of households. To address these concerns, in column (3), we include village-level control variables reflect income levels, which are assets, wealth, liabilities, contributed capital, cumulated savings, and net income. We also include variables that reflect the industry mix, which are the village's income from different sources including cultivation, labor, livestock, fish and shrimp, and business. We find that the estimate is statistically significant and has a similar magnitude. In column (4), we further control for the village's 1-year, 2-year, and 3-year lagged growth rates of assets and wealth, which our model indicates can influence

where branches are opened. The estimate remains statistically significant and has a comparable economic magnitude. A one-minute increase in car travel time from the village to the nearest bank branch is associated with a 1.07% decrease in the village's deposit-cash ratio.

3 Supplemental Information for Model

In this section of the appendix, we provide supplemental information for our model. Subsection 3.1 presents the value functions of the household problem, which we do not describe in Section 3.2 of the main text. Subsection 3.2 derives the equilibrium migration flows and continuation value. Subsection 3.3 sheds light on the model mechanisms by presenting households' choice of occupations, loans, portfolio composition (cash/deposits), and migration. Subsection 3.4 presents the formulas for social welfare. Subsection 3.5 presents the cash flow statements of households. Subsection 3.6 provides a decomposition of GDP growth using three methods of growth accounting: by factor, by credit regime, and by market. Subsection 3.7 discusses the role of market characteristics in determining the locations of branches and further evaluates the model predicted branch locations in different regions and provinces. Subsection 3.8 solves alternative formulations for the choice of branch locations to show the robustness of our results. Finally, subsection 3.9 evaluates how the anticipation and lagged effects of branch openings could affect the DID estimates in the simulated data.

3.1 Value Functions

To complete the households' problem presented in Section 3.2 of the main text, we present the value functions of workers and entrepreneurs (conditional on adjusting the portfolio but not borrowing, or on not adjusting the portfolio but borrowing).

Workers' value function is $W_t(s_t, i) = \max\{W_t^0(s_t, i), W_t^1(s_t, i)\}$. If workers do not adjust the portfolio, their value is $W_t^0(s_t, i)$,

$$W_t^0(s_t, i) = \max_{c_t, m_{t+1}} \left\{ \frac{c_t^{1-\sigma}}{1-\sigma} + \beta \mathbb{E}_t \left[\max_{j \in \{1,\dots,N\}} \{V_{t+1}(s_{t+1}, j) + \epsilon_{j,t+1}\} \right] \right\}, \tag{3.1}$$

$$\text{s.t.} \quad m_{t+1} + c_t = m_t + w_{i,t} - \kappa \mathbb{1}_{\{j \neq i\}},$$

$$a_{t+1} = (1 + r_t) a_t, \quad \text{with} \quad c_t, m_{t+1} \geq 0.$$

If workers adjust the portfolio, their value is $W_t^1(s_t, i)$,

$$W_t^1(s_t,i) = \max_{c_t, m_{t+1}, a_{t+1}} \left\{ \frac{c_t^{1-\sigma}}{1-\sigma} + \beta \mathbb{E}_t \left[\max_{j \in \{1,\dots,N\}} \{V_{t+1}(s_{t+1}, j) + \epsilon_{j,t+1}\} \right] \right\}, \tag{3.2}$$
$$\text{s.t.} \quad m_{t+1} + a_{t+1} + c_t = m_t + (1+r_t)a_t - \zeta_{i,t} + w_{i,t} - \kappa \mathbb{1}_{\{j \neq i\}}, \quad \text{with} \quad c_t, m_{t+1}, a_{t+1} \geq 0.$$

If entrepreneur adjust but do not borrow, their value is $E_t^{10}(s_t, i)$,

$$E_t^{10}(s_t,i) = \max_{c_t, k_t, l_t, m_{t+1}, a_{t+1}} \left\{ \frac{c_t^{1-\sigma}}{1-\sigma} + \beta \mathbb{E}_t \left[\max_{j \in \{1,\dots,N\}} \{V_{t+1}(s_{t+1}, j) + \epsilon_{j,t+1}\} \right] \right\}, \tag{3.3}$$
$$\begin{aligned} \text{s.t.} \quad m_{t+1} + a_{t+1} + c_t =& m_t - k_t + (1+r_t)a_t - \zeta_{i,t} + Z_i z_t (k_t^\alpha l_t^{1-\alpha})^{1-\nu} + (1-\delta)k_t - w_{i,t} l_t - \kappa \mathbb{1}_{\{j \neq i\}}, \\ k_t \leq& m_t, \quad \text{with} \quad c_t, m_{t+1}, a_{t+1} \geq 0. \end{aligned}$$

If entrepreneurs do not adjust but borrow (which requires $m_t \geq \psi_{i,t}$), their value is $E_t^{01}(s_t, i)$,

$$E_t^{01}(s_t,i) = \max_{c_t, k_t, l_t, m_{t+1}} \left\{ \frac{c_t^{1-\sigma}}{1-\sigma} + \beta \mathbb{E}_t \left[\max_{j \in \{1,\dots,N\}} \{V_{t+1}(s_{t+1}, j) + \epsilon_{j,t+1}\} \right] \right\}, \tag{3.4}$$
$$\begin{aligned} \text{s.t.} \quad m_{t+1} + c_t =& -\left[1 + (r_t + \chi)\mathbb{1}_{\{k_t \geq m_t\}}\right][k_t - (m_t - \psi_{i,t})] + Z_i z_t (k_t^\alpha l_t^{1-\alpha})^{1-\nu} \\ &+ (1-\delta)k_t - w_{i,t} l_t - \kappa \mathbb{1}_{\{j \neq i\}}, \\ k_t \leq& (m_t + a_t - \psi_{i,t})/\xi, \\ a_{t+1} =& (1+r_t)a_t, \quad \text{with} \quad c_t, m_{t+1}, a_{t+1} \geq 0. \end{aligned}$$

3.2 Derivation of Migration Flows and Continuation Value

Equilibrium Migration Flows. First, we derive the equilibrium migration flow (equation (3.12) in the main text). The taste shocks $\epsilon_{i,t}$ follow a Type-I extreme value distribution and are independently and identically distributed across time, markets, and households. The cumulative density function $F(\epsilon)$ and probability density function $f(\epsilon)$ are given by

$$F(\epsilon) = \exp\left(-\exp\left(-\frac{\epsilon}{\eta} - \varkappa\right)\right), \tag{3.5}$$
$$f(\epsilon) = \frac{1}{\eta}\exp\left(-\frac{\epsilon}{\eta} - \varkappa - \exp\left(-\frac{\epsilon}{\eta} - \varkappa\right)\right), \tag{3.6}$$

where $\varkappa \equiv \int x \exp(-x - \exp(-x))\mathrm{d}x$ is Euler's constant.

Let $\omega_{i',t}(z_{t+1}|s_t, i)$ denote the proportion of households of type $s_t = \{z_t, m_t, a_t\}$ in market i at t who migrate to market i' and have talent z_{t+1} at $t+1$.[12] As indicated by the term

[12] We distinguish the type of households only by z_{t+1} because the evolution of z_t to z_{t+1} follows a stochastic process. The evolutions of a_t and m_t to a_{t+1} and m_{t+1}, respectively, are deterministic according to the budget

$\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}$ in the households' value function (see equation (3.6) of the main text for example), the migration decisions are based on maximizing utility in period $t+1$. Thus, households who are currently in market i at t choose to migrate to market i' at $t+1$ if and only if $V_{t+1}(z_{t+1},m_{i',t+1},a_{t+1},i')+\epsilon_{i',t+1}\geq V_{t+1}(z_{t+1},m_{j,t+1},a_{t+1},j)+\epsilon_{j,t+1}$ for all $j\neq i'$, where a_{t+1} is given by the optimal policy function of deposits; and $m_{i',t+1}$ and $m_{j,t+1}$ are given by the optimal policy function of cash, which depends on the market (i.e., i' or j) in which households live at $t+1$ because of the fixed migration cost κ paid out of cash (if i' or j is not the same as i). To simplify notations, we write m_{t+1} in the rest of this section, but it should be borne in mind that the value of m_{t+1} also depends on the market in which households live at $t+1$.

We can express $\omega_{i',t}(z_{t+1}|s_t,i)$ as follows:

$$\begin{aligned}&\omega_{i',t}(z_{t+1}|s_t,i)=f(z_{t+1}|z_t)\mathrm{d}z_{t+1}\\&\Pr_t\left[V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\epsilon_{i',t+1}\geq\max_{j\neq i'}\{V_{t+1}(z_{t+1},m_{t+1},a_{t+1},j)+\epsilon_{j,t+1}\}\middle|\,s_t,i,z_{t+1}\right],\end{aligned}\tag{3.7}$$

where the first term $f(z_{t+1}|z_t)\mathrm{d}z_{t+1}$ on the right-hand-side (RHS) of equation (3.7) captures the proportion of households whose talent evolves from z_t to z_{t+1}. The second term, $\Pr_t[\cdots|s_t,i,z_{t+1}]$ captures the proportion of households of type s_t in market i migrating to market i' conditional on z_{t+1}.

We define $\overline{V}_{i'j,t+1}(z_{t+1})\equiv V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')-V_{t+1}(z_{t+1},m_{t+1},a_{t+1},j)$, where our notation omits the dependence of $\overline{V}_{i'j,t+1}(z_{t+1})$ on m_{t+1} and a_{t+1} for brevity because these two variables are deterministic. Equation (3.7) can be simplified as follows:

$$\begin{aligned}&\omega_{i',t}(z_{t+1}|s_t,i)=f(z_{t+1}|z_t)\mathrm{d}z_{t+1}\Pr_t\left[\epsilon_{i',t+1}\geq\max_{j\neq i'}\left\{-\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{j,t+1}\right\}\middle|\,s_t,i,z_{t+1}\right]\\
=&f(z_{t+1}|z_t)\mathrm{d}z_{t+1}\int f(\epsilon_{i',t+1})\Pr_t\left[\epsilon_{i',t+1}\geq\max_{j\neq i'}\left\{-\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{j,t+1}\right\}\middle|\,\epsilon_{i',t+1},s_t,i,z_{t+1}\right]\mathrm{d}\epsilon_{i',t+1}\\
=&f(z_{t+1}|z_t)\mathrm{d}z_{t+1}\int f(\epsilon_{i',t+1})\prod_{j\neq i'}\Pr_t\left[\epsilon_{i',t+1}\geq-\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{j,t+1}\middle|\,\epsilon_{i',t+1},s_t,i,z_{t+1}\right]\mathrm{d}\epsilon_{i',t+1}\\
=&f(z_{t+1}|z_t)\mathrm{d}z_{t+1}\int f(\epsilon_{i',t+1})\prod_{j\neq i'}F(\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{i',t+1})\mathrm{d}\epsilon_{i',t+1}.\end{aligned}\tag{3.8}$$

Substituting equations (3.5) and (3.6) into (3.8), we obtain

$$\begin{aligned}\omega_{i',t}(z_{t+1}|s_t,i)=&f(z_{t+1}|z_t)\mathrm{d}z_{t+1}\int\frac{1}{\eta}\exp\left(-\frac{\epsilon_{i',t+1}}{\eta}-\varkappa\right)\\&\exp\left(-\exp\left(-\frac{\epsilon_{i',t+1}}{\eta}-\varkappa\right)\sum_{j=1}^{N}\exp\left(-\frac{\overline{V}_{i'j,t+1}(z_{t+1})}{\eta}\right)\right)\mathrm{d}\epsilon_{i',t+1}.\end{aligned}\tag{3.9}$$

constraints.

Define

$$\lambda_{t+1}(z_{t+1}, i') = \log\left(\sum_{j=1}^{N} \exp\left(-\frac{\overline{V}_{i'j,t+1}(z_{t+1})}{\eta}\right)\right), \tag{3.10}$$

$$\widetilde{\epsilon}_{i',t+1} = \frac{\epsilon_{i',t+1}}{\eta} + \varkappa - \lambda_{t+1}(z_{t+1}, i'). \tag{3.11}$$

Substituting equations (3.10) and (3.11) into (3.9), we obtain

$$\begin{aligned}\omega_{i',t}(z_{t+1}|s_t, i) =& f(z_{t+1}|z_t)\mathrm{d}z_{t+1} \exp(-\lambda_{t+1}(z_{t+1}, i')) \int \exp(-\widetilde{\epsilon}_{i',t+1} - \exp(-\widetilde{\epsilon}_{i',t+1}))\mathrm{d}\widetilde{\epsilon}_{i',t+1} \\ =& f(z_{t+1}|z_t)\mathrm{d}z_{t+1} \exp(-\lambda_{t+1}(z_{t+1}, i')),\end{aligned} \tag{3.12}$$

where the last equality uses the identity $\int \exp(-x - \exp(-x))\mathrm{d}x = 1$.

Substituting equation (3.10) into (3.12) and using the definition of $\overline{V}_{i'j,t+1}(z_{t+1})$, we obtain

$$\begin{aligned}\omega_{i',t}(z_{t+1}|s_t, i) =& \frac{f(z_{t+1}|z_t)\mathrm{d}z_{t+1}}{\sum_{j=1}^{N} \exp\left(-\frac{\overline{V}_{i'j,t+1}(z_{t+1})}{\eta}\right)} \\ =& f(z_{t+1}|z_t)\mathrm{d}z_{t+1} \frac{\exp\left(V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i')\right)^{1/\eta}}{\sum_{j=1}^{N} \exp(V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j))^{1/\eta}}.\end{aligned} \tag{3.13}$$

Thus, we have

$$\begin{aligned}\omega_{i',t}(s_t, i) =& \int_{z_{t+1}} \omega_{i',t}(z_{t+1}|s_t, i) \\ =& \int_{z_{t+1}} \frac{\exp\left(V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i')\right)^{1/\eta}}{\sum_{j=1}^{N} \exp(V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j))^{1/\eta}} f(z_{t+1}|z_t)\mathrm{d}z_{t+1} \\ =& \mathbb{E}_t\left[\frac{\exp\left(V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i')\right)^{1/\eta}}{\sum_{j=1}^{N} \exp\left(V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j)\right)^{1/\eta}} \middle| s_t, i\right],\end{aligned} \tag{3.14}$$

where $\mathbb{E}_t[\cdot]$ is taken with respect to z_{t+1}.

Continuation Value. Next, we derive the continuation value in the household value function (equation (3.13) of the main text). By the law of iterated expectation, we can rewrite the

continuation value in households' problem as:

$$\begin{aligned}
&\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}\right]\\
\equiv&\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}\middle| s_t,i\right]\\
=&\mathbb{E}_t\left[\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}\middle| s_t,i,z_{t+1}\right]\right],
\end{aligned}\tag{3.15}$$

where the inner $\mathbb{E}_t[\cdot]$ is taken with respect to $\{\epsilon_{j,t+1}\}_{j=1}^N$ conditional on z_{t+1}, and the outer $\mathbb{E}_t[\cdot]$ is taken with respect to z_{t+1}.

Next, following the derivation for equation (3.8), we derive the expression of the inner $\mathbb{E}_t[\cdot]$,

$$\begin{aligned}
&\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}\middle| s_t,i,z_{t+1}\right]\\
=&\sum_{i'=1}^{N}\int(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\epsilon_{i',t+1})\\
&f(\epsilon_{i',t+1})\mathrm{Pr}_t\left[\epsilon_{i',t+1}\geq\max_{j\neq i'}\left\{-\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{j,t+1}\right\}\middle|\epsilon_{i',t+1},s_t,i,z_{t+1}\right]\mathrm{d}\epsilon_{i',t+1}\\
=&\sum_{i'=1}^{N}\int(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\epsilon_{i',t+1})\\
&f(\epsilon_{i',t+1})\prod_{j\neq i'}\mathrm{Pr}_t\left[\epsilon_{i',t+1}\geq-\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{j,t+1}\middle|\epsilon_{i',t+1},s_t,i,z_{t+1}\right]\mathrm{d}\epsilon_{i',t+1}\\
=&\sum_{i'=1}^{N}\int(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\epsilon_{i',t+1})f(\epsilon_{i',t+1})\prod_{j\neq i'}F(\overline{V}_{i'j,t+1}(z_{t+1})+\epsilon_{i',t+1})\mathrm{d}\epsilon_{i',t+1}.
\end{aligned}\tag{3.16}$$

Substituting equations (3.5) and (3.6) into (3.16), we obtain

$$\begin{aligned}
&\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}\middle| s_t,i,z_{t+1}\right]\\
=&\sum_{i'=1}^{N}\int(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\epsilon_{i',t+1})\frac{1}{\eta}\exp\left(-\frac{\epsilon_{i',t+1}}{\eta}-\varkappa\right)\\
&\exp\left(-\exp\left(-\frac{\epsilon_{i',t+1}}{\eta}-\varkappa\right)\sum_{j=1}^{N}\exp\left(-\frac{\overline{V}_{i'j,t+1}(z_{t+1})}{\eta}\right)\right)\mathrm{d}\epsilon_{i',t+1}.
\end{aligned}\tag{3.17}$$

Substituting equations (3.10) and (3.11) into (3.17), we obtain

$$
\begin{aligned}
&\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1},j)+\epsilon_{j,t+1}\}\middle| s_t,i,z_{t+1}\right]\\
=&\sum_{i'=1}^{N}\exp(-\lambda_{t+1}(z_{t+1},i'))\\
&\int\left(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\eta(\widetilde{\epsilon}_{i',t+1}+\lambda_{t+1}(z_{t+1},i')-\varkappa)\right)\exp(-\widetilde{\epsilon}_{i',t+1}-\exp(-\widetilde{\epsilon}_{i',t+1}))\mathrm{d}\widetilde{\epsilon}_{i',t+1}\\
=&\sum_{i'=1}^{N}\exp(-\lambda_{t+1}(z_{t+1},i'))(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\eta\lambda_{t+1}(z_{t+1},i')),
\end{aligned}
\tag{3.18}
$$

where the last equality uses the identities $\int\exp(-x-\exp(-x))\mathrm{d}x=1$ and $\int x\exp(-x-\exp(-x))\mathrm{d}x=\varkappa$.

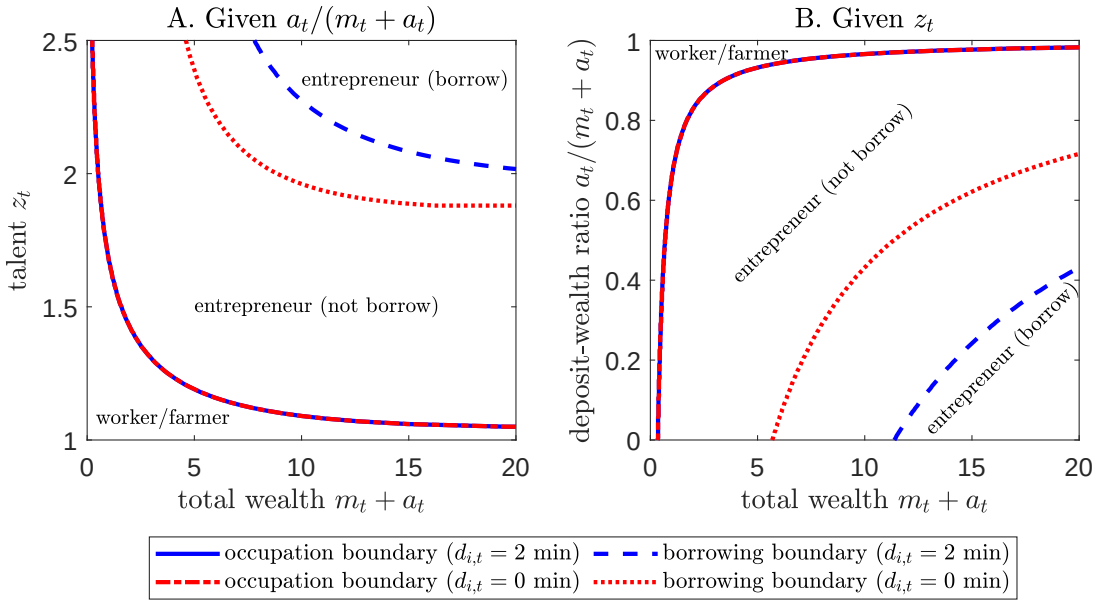
Using equation (3.10), the term $V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\eta\lambda_{t+1}(z_{t+1},i')$ in equation (3.18) can be written as

$$
\begin{aligned}
&V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\eta\lambda_{t+1}(z_{t+1},i')\\
=&V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\eta\log\left(\sum_{j=1}^{N}\exp\left(-\frac{\overline{V}_{i'j,t+1}(z_{t+1})}{\eta}\right)\right)\\
=&V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')+\eta\log\left(\sum_{j=1}^{N}\exp\left(\frac{V_{t+1}(z_{t+1},m_{t+1},a_{t+1},j)-V_{t+1}(z_{t+1},m_{t+1},a_{t+1},i')}{\eta}\right)\right)\\
=&\eta\log\left(\sum_{j=1}^{N}\exp\left(\frac{V_{t+1}(z_{t+1},m_{t+1},a_{t+1},j)}{\eta}\right)\right).
\end{aligned}
\tag{3.19}
$$

Note that the RHS of equation (3.19) does not depend on i'. We substitute equation (3.19) into (3.18) and obtain,

$$
\begin{aligned}
&\mathbb{E}_t\left[\max_{j\in\{1,\ldots,N\}}\{V_{t+1}(s_{t+1})+\epsilon_{j,t+1}\}\middle| s_t,i,z_{t+1}\right]\\
=&\left(\sum_{i'=1}^{N}\exp(-\lambda_{t+1}(z_{t+1},i'))\right)\left(\eta\log\left(\sum_{j=1}^{N}\exp\left(\frac{V_{t+1}(z_{t+1},m_{t+1},a_{t+1},j)}{\eta}\right)\right)\right)\\
=&\eta\log\left(\sum_{j=1}^{N}\exp(V_{t+1}(z_{t+1},m_{t+1},a_{t+1},j))^{1/\eta}\right),
\end{aligned}
\tag{3.20}
$$

where the last equality uses $\sum_{i'=1}^{N}\exp(-\lambda_{t+1}(z_{t+1},i'))=1$ by substituting out $\lambda_{t+1}(z_{t+1},i')$ using (3.10). Finally, substituting equation (3.20) into (3.15), we derive equation (3.13) of the main text.



Note: This figure illustrates occupation choices and borrowing decisions for households with different levels of talent z_t, cash m_t, and deposits a_t, living in the same market i. The x-axis is total wealth $m_t + a_t$ for both panels. In panel A, we set the deposit-wealth ratio at $a_t/(m_t + a_t) = 0.2$, corresponding to the median value across markets, and vary talent z_t along the y-axis. In panel B, we set $z_t = 2.1$, corresponding to the median value of entrepreneurs with credit access, and vary $a_t/(m_t + a_t)$ along the y-axis. The blue solid and dashed lines represent the boundaries if $d_{i,t} = 2$ minutes. The red dash-dotted and dotted lines represent the boundaries if $d_{i,t} = 0$. We focus on a partial-equilibrium setting with a fixed interest rate $r_t = 8.3\%$ and local wage $w_{i,t} = 0.77$, corresponding to the median wage across markets in 1986. We set $\zeta_{i,t} = \kappa = \infty$ so that portfolio adjustment and migration decisions are irrelevant. Other parameters are set according to our calibration in Table 2 of the main text.

Figure OA.5: Occupation and borrowing decisions.

3.3 Illustration of Model Mechanisms

Our model incorporates multiple markets to focus on the analysis of a spatial equilibrium. Dynamic bank expansion reduces the distance from nearest branch $d_{i,t}$ across markets, leading to lower credit entry costs $\psi_{i,t}$ and portfolio adjustment costs $\zeta_{i,t}$. This in turn affects the distribution of loans and deposits across markets, consequently influencing households' income, occupation choice, and migration decisions. In this section of the appendix, we elucidate the model mechanisms by examining the households' choice of occupations, loans, portfolio composition (cash/deposits), and migration.

3.3.1 Occupation and Borrowing Decisions

Consider households of type s_t in market i. Panel A of Figure OA.5 depicts the occupation and borrowing decisions of households with different values of $m_t + a_t$ (x-axis) and z_t (y-axis), given a

fixed $a_t/(m_t + a_t)$.[13] Panel B complements panel A by representing $m_t + a_t$ along the x-axis and $a_t/(m_t + a_t)$ along the y-axis, taking z_t as fixed. Households' occupation and borrowing decisions can be sufficiently summarized by the occupation boundary and the borrowing boundary.

First, consider the case where $d_{i,t} = 2$ minutes. In panel A (B) of Figure OA.5, households to the southwest (northwest) of the occupation boundary (the blue solid line) choose to be workers or farmers because they do not have sufficient wealth to start a profitable business, whereas households to the northeast (southeast) choose to be entrepreneurs. In panel A, the occupation boundary is downward sloping, indicating that more talented households require less wealth to become entrepreneurs because their businesses can be profitable at smaller scales owing to higher productivity. In panel B, the occupation boundary is upward sloping, implying that households with a higher deposit-wealth ratio need more total wealth to start a profitable business owing to the illiquidity of deposits (i.e., buying capital requires cash).

The blue dashed lines in each panel indicate the borrowing boundaries, which are to the right of the occupation boundaries because only entrepreneurs can obtain bank loans. Entrepreneurs in the region enclosed by the borrowing boundary pay the credit entry cost $\psi_{i,t}$ to borrow. Entrepreneurs in the regions outside the borrowing boundaries do not borrow, either because they do not have enough cash to pay $\psi_{i,t}$ or because their benefit from loans is less than the credit entry cost.

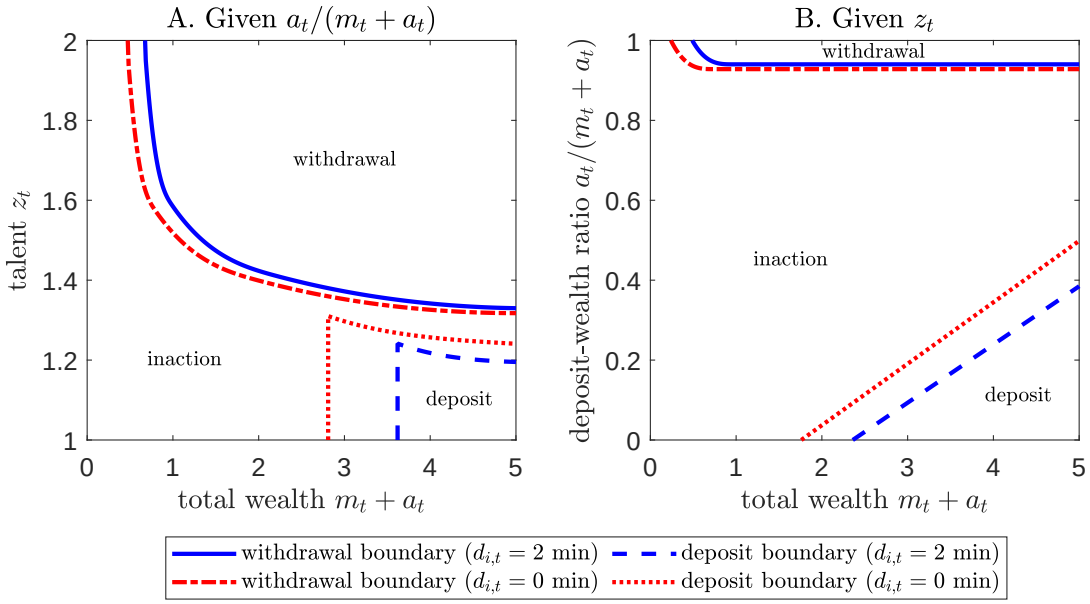
Moving from the case where $d_{i,t} = 2$ minutes to that where $d_{i,t} = 0$, we observe that the occupation boundary remains unchanged, but the borrowing boundaries in both panels shift to the left (from the blue dashed lines to the red dotted lines), implying that more entrepreneurs obtain loans when the credit entry cost $\psi_{i,t}$ is lower. Thus, when all else is equal, bank expansion raises the credit access ratio by reducing $\psi_{i,t}$ through a lower $d_{i,t}$.

3.3.2 Portfolio Adjustment Decisions

While the credit entry cost $\psi_{i,t}$ largely determines households' occupation and borrowing decisions, the portfolio adjustment cost $\zeta_{i,t}$ mostly determines their portfolio adjustment decisions. Figure OA.6 illustrates households' withdrawal and deposit boundaries.

Focusing on the case where $d_{i,t} = 2$ minutes, households above the withdrawal boundary (shown by the blue solid lines in panels A and B) withdraw funds, i.e., convert a fraction of their deposits to cash. Intuitively, in panel A, given the same total wealth $m_t + a_t$, households with higher talent z_t operate businesses with higher returns, which motivates them to withdraw funds to buy more capital. The withdrawal boundary is downward sloping because more talented households are more eager to withdraw at lower wealth levels due to the higher returns from their

[13] Households in market i are characterized by $s_t = \{z_t, m_t, a_t\}$. To explain the intuition, we change variables and focus on talent z_t, total wealth $m_t + a_t$, and the deposit-wealth ratio $a_t/(m_t + a_t)$.



Note: This figure illustrates the portfolio adjustment decisions for households with different levels of talent z_t, cash m_t, and deposits a_t, living in the same market i. The x-axis is total wealth $m_t + a_t$ for both panels. In panel A, we set the deposit-wealth ratio at $a_t/(m_t + a_t) = 0.2$, corresponding to the median value across markets, and vary talent z_t along the y-axis. In panel B, we set $z_t = 1.16$, corresponding to the median value of households, and vary $a_t/(m_t + a_t)$ along the y-axis. The blue solid and dashed lines represent the boundaries if $d_{i,t} = 2$ minutes. The red dash-dotted and dotted lines represent the boundaries if $d_{i,t} = 0$. We focus on a partial-equilibrium setting with a fixed interest rate $r_t = 8.3\%$ and local wage $w_{i,t} = 0.77$, corresponding to the median wage across markets in 1986. We set $\psi_{i,t} = \kappa = \infty$ so that borrowing and migration decisions are irrelevant. Other parameters are set according to our calibration in Table 2 of the main text.

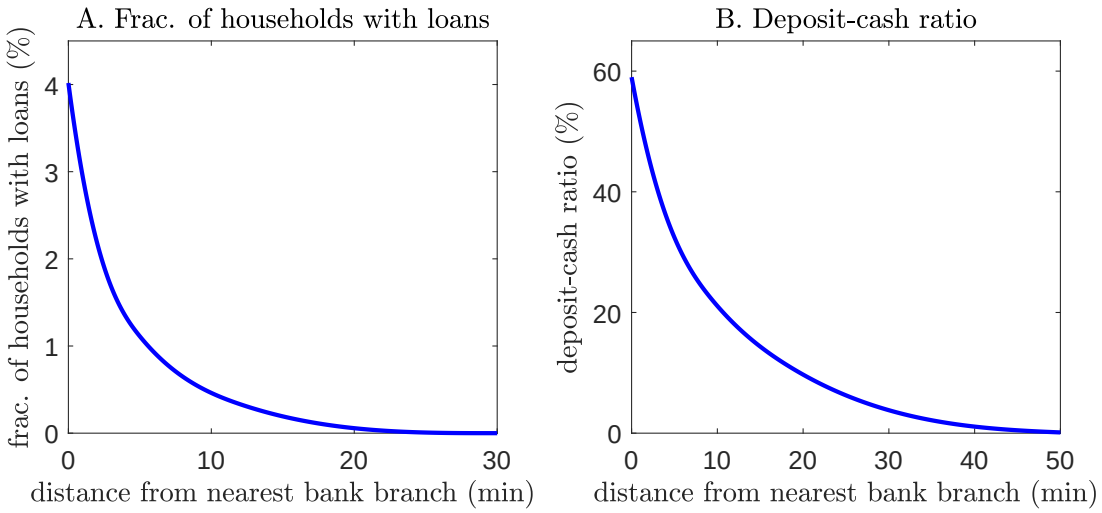
Figure OA.6: Portfolio adjustment decisions.

businesses. In panel B, given $m_t + a_t$, households with high deposit-wealth ratios withdraw funds because they need cash to buy consumption goods or capital.[14]

The blue dashed line in each panel indicates the deposit boundary. Households in the lower-right region enclosed by the deposit boundary make a deposit (i.e., convert a fraction of their cash to deposits). These households have high wealth and hold much more cash than they need, either because their talent is low (panel A) or because their deposit-wealth ratio is low (panel B). Thus, they choose to put some of their cash into the bank to earn interest.

The remaining region indicates inaction. Households in this region keep their current portfolio of cash and deposits unchanged to save the portfolio adjustment cost $\zeta_{i,t}$. Moving from the case where $d_{i,t} = 2$ minutes to that where $d_{i,t} = 0$, we observe that both the withdrawal and deposit regions expand and the inaction region shrinks, as households choose to optimize their portfolios more frequently when the cost $\zeta_{i,t}$ of doing so falls through a lower $d_{i,t}$.

[14]When total wealth $m_t + a_t$ is very low, households do not withdraw even when the deposit-wealth ratio is very high, because the benefit of having more cash is outweighed by the portfolio adjustment cost $\zeta_{i,t}$.



Note: This figure plots the model-implied fraction of households with bank loans (panel A) and the average deposit-cash ratio of households (panel B) as a function of the distance from the nearest bank branch $d_{i,t}$. We use the initial distribution of households in 1986 when computing the average values.

Figure OA.7: Model-implied relationship between the fraction of households with bank loans, the average deposit-cash ratio, and the distance from the nearest bank branch.

3.3.3 Financial Decisions and Distances from a Branch

Panel A of Figure OA.7 plots the model-implied relationship between the fraction of households with bank loans and the distance from the nearest bank branch $d_{i,t}$, based on the initial distribution of households in 1986. The fraction of households with bank loans is approximately 4.0% when $d_{i,t} = 0$, and it decreases to 0% when $d_{i,t} = 30$ minutes. Panel B of Figure OA.7 plots the relationship between the average deposit-cash ratio of households and the distance from the nearest bank branch $d_{i,t}$. The average deposit-cash ratio drops from approximately 59% to 0% when $d_{i,t}$ increases from 0 to 50 minutes.

3.3.4 Migration Decisions

We now illustrate households' migration decisions. Idiosyncratic taste shocks simply introduce randomness into migration flows across markets. For clarity, we remove idiosyncratic taste shocks by setting $\eta = 0$ in this section, and we discuss the impacts of idiosyncratic taste shocks in Online Appendix 4.1.1.

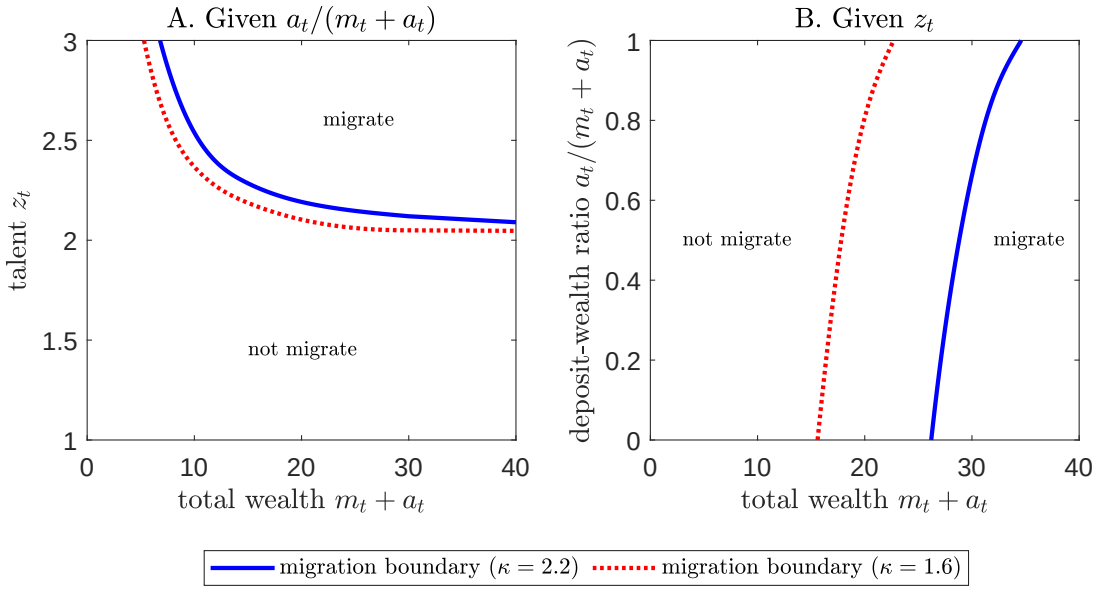
Households' migration decisions crucially depend on the market in which they live. In our model, markets differ from each other in three aspects, market size $\Pi_{i,t}$, market productivity Z_i, and the market's distance from the nearest bank branch $d_{i,t}$. Market size $\Pi_{i,t}$ is a scale factor that determines a market's contribution to the whole economy, so it has little impact on households' migration decisions. The impact of market productivity Z_i on migration decisions is straight

forward: all else being equal, households prefer to migrate from markets with low productivity to those with high productivity. The market's distance from the nearest bank branch $d_{i,t}$ has nontrivial impacts on migration decisions because $d_{i,t}$ can interact with households' talent z_t, cash holdings m_t and deposits a_t to determine their occupation choice and borrowing decisions. Below, we illustrate households' migration decisions in two representative markets, one located far from bank branches and one with a local bank branch.

Market without a Branch. First, consider households of type s_t in market i without a branch. Panel A of Figure OA.8 illustrates the migration decisions of households with different values of $m_t + a_t$ (x-axis) and z_t (y-axis), given that $a_t/(m_t + a_t)$ is fixed. Households to the northeast of the migration boundary in panel A of Figure OA.8 choose to migrate to markets close to bank branches to lower their credit access costs and earn higher business incomes. Most of these households are entrepreneurs with a high level of talent z_t. At the same time, these households' businesses are constrained by their wealth and thus, they have to borrow from the bank. Migrating to markets close to bank branches allows them to borrow at lower credit entry costs. The migration boundary is downward sloping because the collateral constraint restricts the amount of bank loans that households can obtain. For households with less total wealth $m_t + a_t$, there will be less remaining wealth after paying the pecuniary migration cost κ, and thus households cannot borrow much after migration. As a result, only households with higher talent levels z_t will choose to migrate because the marginal value of borrowing is higher for these households.

Panel B of Figure OA.8 complements panel A by replacing the level of talent on the y-axis with the deposit-wealth ratio $a_t/(m_t + a_t)$, assuming that talent is fixed at $z_t = 2.1$, which corresponds to the median value of entrepreneurs with credit access. Households to the right of the migration boundary choose to migrate. These households are constrained entrepreneurs, similar to the households to the northeast of the migration boundary in panel A of Figure OA.8. Migrating to markets close to bank branches allows them to borrow at lower credit entry costs than in their original place of residence. Households to the left of the migration boundary do not migrate because the migration cost κ outweighs the benefit of obtaining loans after migration; these households cannot obtain large loans due to their low wealth and collateral constraints. The migration boundary in panel B of Figure OA.8 is upward sloping because as the deposit-wealth ratio $a_t/(m_t + a_t)$ increases, households will have less wealth in cash m_t in the current period, which restricts the amount of capital they can purchase. Because less capital is used for production, households' current business income is lower than it would be otherwise, implying that households need to have more total wealth $m_t + a_t$ at the beginning of period t to achieve the threshold level of wealth above which they would choose to migrate in the next period.

Moving from the case where $\kappa = 2.2$ to that where $\kappa = 1.6$, we observe that the migration boundaries in both panels of Figure OA.8 shift to the left (from blue solid lines to red dotted

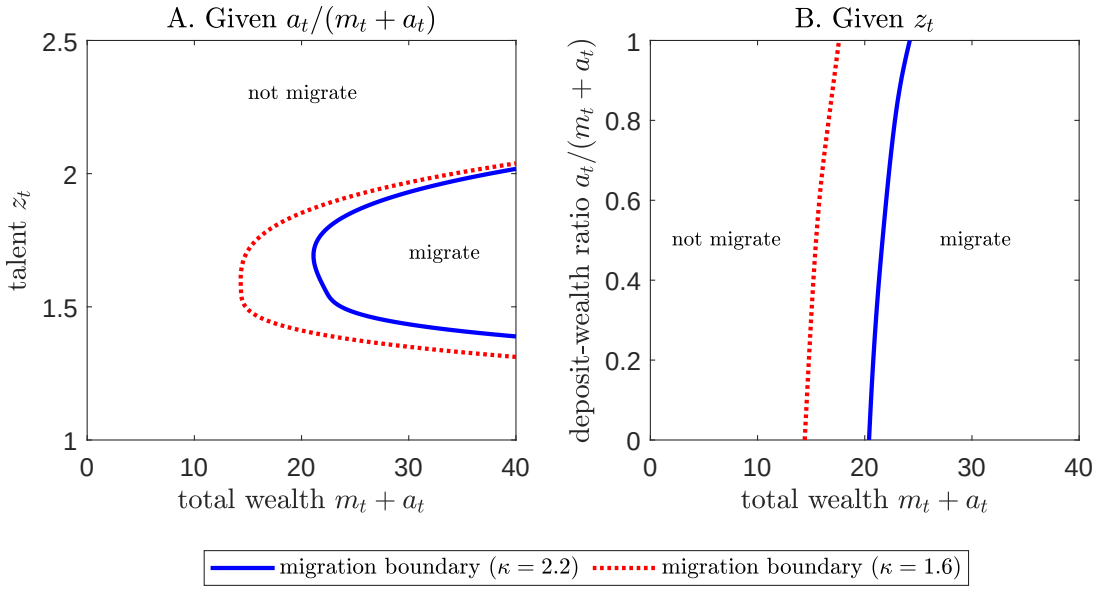


Note: This figure illustrates migration decisions for households with different levels of talent z_t, cash m_t, and deposits a_t, living in a market that is far away from bank branches, with $d_{i,t} = 12.8$ minutes, which corresponds to the median car travel time to the nearest bank branch among all markets without branches in 1986. The x-axis is total wealth $m_t + a_t$ for both panels. In panel A, we set the deposit-wealth ratio at $a_t/(m_t + a_t) = 0.2$, corresponding to the median value across markets, and vary talent z_t along the y-axis. In panel B, we set $z_t = 2.1$, corresponding to the median value of entrepreneurs with credit access, and vary $a_t/(m_t + a_t)$ along the y-axis. We focus on a partial-equilibrium setting with $r_t = 8.3\%$ and $\{w_{i,t}\}_{i=1}^{N}$ according to the equilibrium interest rate and wages in 1986. The blue solid lines represent the case with $\kappa = 2.2$ according to our calibration in Table 2 of the main text. The red dotted lines represent the case with a lower fixed migration cost, $\kappa = 1.6$. We set $\eta = 0$ to remove idiosyncratic taste shocks. Other parameters are set according to our calibration in Table 2 of the main text.

Figure OA.8: Migration decisions of households living far away from bank branches.

lines) and the region where households choose to migrate expands. More households will migrate when the fixed migration cost κ is lower.

Market with a Branch. Now, consider households of type s_t in market i with a branch. Panel A of Figure OA.9 shows that households to the right of the migration boundary choose to migrate to markets far away from bank branches. It may seem counterintuitive that some households prefer to live far away from bank branches. In fact, these households are wealthy entrepreneurs with only moderate levels of talent. They do not operate their businesses on a large scale because of their relatively low talent. Thus, living in markets with bank branches does not really benefit these entrepreneurs because they do not borrow. However, the relatively high local wages in markets with bank branches imply high costs of production. Thus, these entrepreneurs of mediocre talent have the incentive to move to a distant market far from bank branches, as they can hire workers in this market at low local wages. In equilibrium, entrepreneurs who want to migrate to markets



Note: This figure illustrates migration decisions for households with different levels of talent z_t, cash m_t, and deposits a_t, living in a market with a branch, with $d_{i,t} = 0$. The x-axis is total wealth $m_t + a_t$ for both panels. In panel A, we set the deposit-wealth ratio at $a_t/(m_t + a_t) = 0.2$, corresponding to the median value across markets, and vary talent z_t along the y-axis. In panel B, we set $z_t = 1.75$, corresponding to the median value of entrepreneurs, and vary $a_t/(m_t + a_t)$ along the y-axis. We focus on a partial-equilibrium setting with $r_t = 8.3\%$ and $\{w_{i,t}\}_{i=1}^{N}$ according to the equilibrium interest rate and wages in 1986. The blue solid lines represent the case with $\kappa = 2.2$ according to our calibration in Table 2 of the main text. The red dotted lines represent the case with a lower fixed migration cost, $\kappa = 1.6$. We set $\eta = 0$ to remove idiosyncratic taste shocks. Other parameters are set according to our calibration in Table 2 of the main text.

Figure OA.9: Migration decisions of households living in markets with a branch.

far away from bank branches only account for a tiny fraction of the migration flow when $\eta = 0$. Thus, while theoretically possible, the migration decisions described in panel A of Figure OA.9 are not quantitatively important. However, once we increase the value of η, more households will migrate for idiosyncratic reasons, and hence, a larger migration flow out of markets with bank branches will be observed in equilibrium. By adjusting the parameter η, we are able to match the two-way migration flows across markets with and without branches observed in the data (see Online Appendix 4.1.1 for details).

The shape of the migration boundary in panel A of Figure OA.9 can be explained as follows. When deciding whether to migrate to markets without bank branches, entrepreneurs face a tradeoff between borrowing in their current markets at lower costs or hiring workers at lower wages in the more distant markets to which they could migrate. As z_t increases, entrepreneurs' migration decisions are affected through two channels. First, a higher z_t increases entrepreneurs' demand for loans because they want to operate their businesses on a larger scale; this motivates them to remain in their current market rather than migrating to markets far away from bank

branches. Second, a higher z_t increases entrepreneurs' demand for labor, relative to their demand for capital, so wage costs account for a larger fraction of their overall production cost. This change in the composition of production costs motivates entrepreneurs to migrate to markets far away from bank branches to benefit from lower local wages. The second channel dominates the first only for entrepreneurs with low z_t, as these entrepreneurs have a low demand for loans in any case. For $z_t \geq 1.7$, the blue solid line is upward sloping because the first channel dominates the second. As z_t increases, entrepreneurs are increasingly tempted to increase their scale of business by borrowing. Thus, if they chose to migrate to markets without bank branches, they would require more wealth (so that they can avoid borrowing after migration). For $z_t \leq 1.7$, the blue solid line is downward sloping because the second channel dominates the first. As z_t increases, wage costs account for a larger fraction of the overall production cost. Thus, even less wealthy entrepreneurs who do not operate businesses on a large scale are motivated to migrate to markets far away from branches, in pursuit of lower local wages.

Panel B of Figure OA.9 complements panel A by replacing talent on the y-axis with the deposit-wealth ratio $a_t/(m_t + a_t)$, assuming that talent is fixed at $z_t = 1.75$, corresponding to the median value of entrepreneurs. Households to the right of the migration boundary choose to migrate to benefit from lower local wages in destination markets than in their original market, as explained in relation to panel A of this figure.

Moving from the case where $\kappa = 2.2$ to that where $\kappa = 1.6$, we observe that the migration boundaries in both panels of Figure OA.8 shift to the left (from blue solid lines to red dotted lines), implying that more households will migrate when the fixed migration cost κ is lower.

3.4 Welfare Calculation

Following Jones and Klenow (2016), we use our model to compute welfare differences based on a consumption-equivalent measure over time (Section 6.3 of the main text).

Consider a household of type $s = \{z, m, a\}$ in market i at $t = 0$. The household's expected lifetime utility is

$$V_0(s, i) = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma} \right], \tag{3.21}$$

where $\{c_t\}_{t=0}^{\infty}$ represents her consumption at $t \geq 0$.

Now, consider another household of the same type s in market i but at a different time $t = \tau > 0$. The expected lifetime utility is

$$V_\tau(s, i) = \mathbb{E}_\tau \left[\sum_{t=\tau}^{\infty} \beta^{t-\tau} \frac{(c'_t)^{1-\sigma}}{1-\sigma} \right], \tag{3.22}$$

where $\{c'_t\}_{t=\tau}^{\infty}$ represents her consumption at $t \geq \tau$.

Denote by $\lambda_{0\to\tau}(s,i)$ the consumption-equivalent measure of welfare change for households of type s in market i over the period from $t=0$ to $t=\tau$. We compute $\lambda_{0\to\tau}(s,i)$ as follows:

$$\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\frac{(\lambda_{0\to\tau}(s,i)c_t)^{1-\sigma}}{1-\sigma}\right]=\mathbb{E}_\tau\left[\sum_{t=\tau}^{\infty}\beta^{t-\tau}\frac{(c'_t)^{1-\sigma}}{1-\sigma}\right]. \tag{3.23}$$

Intuitively, $\lambda_{0\to\tau}(s,i)$ is a factor that augments the consumption path of the household at $t=0$ so that she has the same expected lifetime utility as the household at $t=\tau$. This is a relevant comparison because both households have the same type s and live in the same market i.

Because the utility function is homogeneous of degree $1-\sigma$, we have

$$\lambda_{0\to\tau}(s,i)^{1-\sigma}\mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\frac{c_t^{1-\sigma}}{1-\sigma}\right]=\mathbb{E}_\tau\left[\sum_{t=\tau}^{\infty}\beta^{t-\tau}\frac{(c'_t)^{1-\sigma}}{1-\sigma}\right], \tag{3.24}$$

which gives

$$\lambda_{0\to\tau}(s,i)=\left[\frac{V_\tau(s,i)}{V_0(s,i)}\right]^{\frac{1}{1-\sigma}}. \tag{3.25}$$

Equation (3.25) derives the welfare change of type-s households in market i. The market-wide welfare change is computed as the average welfare change of all households in market i (ignoring the change in market size due to migration) using a utilitarian social welfare function:

$$\lambda_{i,0\to\tau}=\left[\frac{\int V_\tau(s,i)\phi_\tau(s|i)\mathrm{d}s}{\int V_0(s,i)\phi_0(s|i)\mathrm{d}s}\right]^{\frac{1}{1-\sigma}}. \tag{3.26}$$

The economy-wide welfare change is obtained by aggregating the welfare change of all households living in every market:

$$\lambda_{0\to\tau}=\left[\frac{\sum_{i=1}^{N}\Pi_{i,t}\int V_\tau(s,i)\phi_\tau(s|i)\mathrm{d}s}{\sum_{i=1}^{N}\Pi_{i,t}\int V_0(s,i)\phi_0(s|i)\mathrm{d}s}\right]^{\frac{1}{1-\sigma}}. \tag{3.27}$$

3.5 Cash Flow Statements

Tables OA.9 and OA.10 present the cash flow statements of workers (and farmers) and entrepreneurs in market i.

Table OA.9: The cash flow statement of workers (and farmers).

(1) **Initial cash holdings at the beginning of period** t	m_t
(2) *Wage (or subsistence income)*	$w_{i,t}$ (or fZ_i)
Pay portfolio adjustment cost	$-\zeta_{i,t}$
Invest in deposits	$(1+r_t)a_t - a_{t+1}$
(3) *Cash from portfolio adjustment*	$(1+r_t)a_t - a_{t+1} - \zeta_{i,t}$
(4) *Consumption cash flows*	$-c_t$
(5) *Cost of migration*	$-\kappa$
(6) **Cash holdings at the end of period** t	$m_{t+1} =$
Case 1: no adjustment, no migration: (1)+(2)+(4)	$m_t + w_{i,t} - c_t$
Case 2: adjustment, no migration: (1)+(2)+(3)+(4)	$m_t + w_{i,t} - c_t + (1+r_t)a_t - a_{t+1} - \zeta_{i,t}$
Case 3: no adjustment, migration: (1)+(2)+(4)+(5)	$m_t + w_{i,t} - c_t - \kappa$
Case 4: adjustment, migration: (1)+(2)+(3)+(4)+(5)	$m_t + w_{i,t} - c_t + (1+r_t)a_t - a_{t+1} - \zeta_{i,t} - \kappa$

3.6 Decomposition of Aggregate Growth

In this appendix section, we provide a decomposition of GDP growth using standard formulas of growth accounting. We present the formulas in Subsection 3.6.1 and results in Subsection 3.6.2

3.6.1 Formulas

The economy's aggregate output (GDP) Y_t and TFP Z_t are defined as

$$Y_t = \sum_{i=1}^{N} \left[\Pi_{i,t} \int Z_i z \left(k_t(s,i)^{\alpha} l_t(s,i)^{1-\alpha} \right)^{1-\nu} \phi_t(s|i) \mathrm{d}s \right], \tag{3.28}$$

$$Z_t = Y_t / (K_t^{\alpha} L_t^{1-\alpha}), \tag{3.29}$$

where K_t and L_t are aggregate capital and labor, respectively, given by

$$K_t = \sum_{i=1}^{N} \left[\Pi_{i,t} \int k_t(s,i) \phi_t(s|i) \mathrm{d}s \right], \tag{3.30}$$

$$L_t = \sum_{i=1}^{N} \left[\Pi_{i,t} \int l_t(s,i) \phi_t(s|i) \mathrm{d}s \right]. \tag{3.31}$$

Growth Accounting by Factor. According to equation (3.29),

$$Y_t = Z_t K_t^{\alpha} L_t^{1-\alpha}, \tag{3.32}$$

Table OA.10: The cash flow statement of entrepreneurs.

(1)	**Initial cash holdings at the beginning of period t**	m_t
	Cash flows at the beginning of period t (before production)	
	Pay credit entry cost	$-\psi_{i,t}$
	Borrow loans at the beginning of period t	$k_t - (m_t - \psi_{i,t})$
(2)	*Cash from financing (conditional on financing)*	$k_t - (m_t - \psi_{i,t}) - \psi_{i,t}$
	Buy working capital from capital shop	$-k_t$
(3)	*Cash from operations*	$-k_t$
	Cash flows at the end of period t (after production)	
	Repay at the end of period t	$-(1 + r_t + \chi)\left[k_t - (m_t - \psi_{i,t})\right]$
(4)	*Cash from financing (conditional on financing)*	$-(1 + r_t + \chi)\left[k_t - (m_t - \psi_{i,t})\right]$
	Sell goods to households	$Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu}$
	Pay wages to workers	$-w_{i,t} l_t$
	Sell undepreciated working capital to capital shop	$(1-\delta)k_t$
(5)	*Cash from operations*	$Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} + (1-\delta)k_t - w_{i,t} l_t$
	Pay portfolio adjustment cost	$-\zeta_{i,t}$
	Invest in deposits	$(1 + r_t)a_t - a_{t+1}$
(6)	*Cash from portfolio adjustment*	$(1 + r_t)a_t - a_{t+1} - \zeta_{i,t}$
(7)	*Consumption cash flows*	$-c_t$
(8)	*Cost of migration*	$-\kappa$
(9)	**Cash holdings at the end of period t**	$m_{t+1} =$
	Case 1: no financing, no adjustment, no migration: (1)+(3)+(5)+(7)	$m_t + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t$
	Case 2: no financing, no adjustment, migration: (1)+(3)+(5)+(7)+(8)	$m_t + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t - \kappa$
	Case 3: no financing, adjustment, no migration: (1)+(3)+(5)+(6)+(7)	$m_t + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t + (1 + r_t)a_t - a_{t+1} - \zeta_{i,t}$
	Case 4: no financing, adjustment, migration: (1)+(3)+(5)+(6)+(7)+(8)	$m_t + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t + (1 + r_t)a_t - a_{t+1} - \zeta_{i,t} - \kappa$
	Case 5: financing, no adjustment, no migration: (1)+(2)+(3)+(4)+(5)+(7)	$m_t - (r_t + \chi)\left[k_t - (m_t - \psi_{i,t})\right] - \psi_{i,t} + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t$
	Case 6: financing, no adjustment, migration: (1)+(2)+(3)+(4)+(5)+(7)+(8)	$m_t - (r_t + \chi)\left[k_t - (m_t - \psi_{i,t})\right] - \psi_{i,t} + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t - \kappa$
	Case 7: financing, adjustment, no migration: (1)+(2)+(3)+(4)+(5)+(6)+(7)	$m_t - (r_t + \chi)\left[k_t - (m_t - \psi_{i,t})\right] - \psi_{i,t} + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t + (1 + r_t)a_t - a_{t+1} - \zeta_{i,t}$
	Case 8: financing, adjustment, migration: (1)+(2)+(3)+(4)+(5)+(6)+(7)+(8)	$m_t - (r_t + \chi)\left[k_t - (m_t - \psi_{i,t})\right] - \psi_{i,t} + Z_i z_t (k_t^{\alpha} l_t^{1-\alpha})^{1-\nu} - \delta k_t - w_{i,t} l_t - c_t + (1 + r_t)a_t - a_{t+1} - \zeta_{i,t} - \kappa$

where Y_t, Z_t, K_t, and L_t are the economy's aggregate output, TFP, capital, and labor, respectively, as defined in equations (3.28) to (3.31). Taking the logarithm of both sides of equation (3.32),

$$\log(Y_t) = \log(Z_t) + \alpha \log(K_t) + (1-\alpha)\log(L_t). \tag{3.33}$$

Moving from t to $t+1$, we obtain a similar expression:

$$\log(Y_{t+1}) = \log(Z_{t+1}) + \alpha \log(K_{t+1}) + (1-\alpha)\log(L_{t+1}). \tag{3.34}$$

Taking the difference of $\log(Y_t)$ and $\log(Y_{t+1})$ using equations (3.33) and (3.34) yields

$$\log\left(\frac{Y_{t+1}}{Y_t}\right) = \log\left(\frac{Z_{t+1}}{Z_t}\right) + \alpha \log\left(\frac{K_{t+1}}{K_t}\right) + (1-\alpha)\log\left(\frac{L_{t+1}}{L_t}\right). \tag{3.35}$$

Equation (3.35) can be approximated by the first difference:

$$\frac{Y_{t+1}-Y_t}{Y_t} = \frac{Z_{t+1}-Z_t}{Z_t} + \alpha\frac{K_{t+1}-K_t}{K_t} + (1-\alpha)\frac{L_{t+1}-L_t}{L_t}. \tag{3.36}$$

Let $g_t^x = (x_{t+1}-x_t)/x_t$ denote the percentage change in the value of variable x from time t to $t+1$. Thus equation (3.36) can be written as

$$g_t^Y = \underbrace{g_t^Z}_{\text{TFP growth}} + \underbrace{\alpha g_t^K}_{\text{capital growth}} + \underbrace{(1-\alpha)g_t^L}_{\text{labor growth}}. \tag{3.37}$$

Growth Accounting by Credit Regime. We conduct growth accounting by credit regime following Dabla-Norris et al. (2021). Denote by $y_t(s,i)$ the output of an entrepreneur of type $s = \{z, m, a\}$ in market i at time t. Denote by $\Omega_{i,t}^c$ the set that includes all types of entrepreneurs with credit access (i.e., those in the credit regime), and by $\Omega_{i,t}^n$ the set that includes all types of entrepreneurs without credit access in market i at time t (i.e., those in the non-credit regime). That is,

$$\Omega_{i,t}^c = \left\{ s : E_t(s,i) \geq \max\{F_t(s,i), W_t(s,i)\}, \max\{E_t^{01}(s,i), E_t^{11}(s,i)\} \geq \max\{E_{i,t}^{00}(s,i), E_{i,t}^{10}(s,i)\} \right\}; \tag{3.38}$$

$$\Omega_{i,t}^n = \left\{ s : E_t(s,i) \geq \max\{F_t(s,i), W_t(s,i)\}, \max\{E_t^{01}(s,i), E_t^{11}(s,i)\} < \max\{E_t^{00}(s,i), E_t^{10}(s,i)\} \right\}. \tag{3.39}$$

The economy's GDP at time t can be written as

$$Y_t = \sum_{i=1}^{N} \Pi_{i,t} \left[\int_{\Omega^c_{i,t}} y_t(s,i)\phi_t(s|i)\mathrm{d}s + \int_{\Omega^n_{i,t}} y_t(s,i)\phi_t(s|i)\mathrm{d}s \right]. \tag{3.40}$$

Thus,

$$\begin{aligned} Y_{t+1} - Y_t =& \sum_{i=1}^{N} \Pi_{i,t+1} \left[\int_{\Omega^c_{i,t+1}} y_{t+1}(s,i)\phi_{t+1}(s|i)\mathrm{d}s + \int_{\Omega^n_{i,t+1}} y_{t+1}(s,i)\phi_{t+1}(s|i)\mathrm{d}s \right] \\ & - \sum_{i=1}^{N} \Pi_{i,t} \left[\int_{\Omega^c_{i,t}} y_t(s,i)\phi_t(s|i)\mathrm{d}s + \int_{\Omega^n_{i,t}} y_t(s,i)\phi_t(s|i)\mathrm{d}s \right]. \end{aligned} \tag{3.41}$$

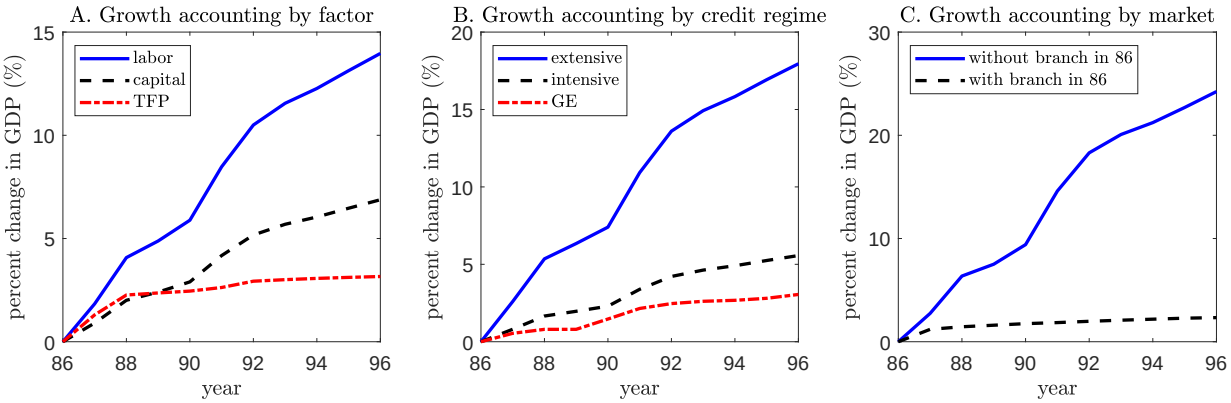
For each market i, let $I_{i,t} = \Omega^c_{i,t} \cap \Omega^c_{i,t+1}$ be the set that includes all types of entrepreneurs who have credit access at both t and $t+1$. Let $E_{i,t} = \left(\Omega^n_{i,t} \cap \Omega^c_{i,t+1}\right) \cup \left(\Omega^c_{i,t} \cap \Omega^n_{i,t+1}\right)$ be the set that includes all types of entrepreneurs whose credit access conditions are changed from t to $t+1$. Finally, let $G_{i,t} = \Omega^n_{i,t} \cap \Omega^n_{i,t+1}$ be the set that includes all types of entrepreneurs who do not have credit access at both t and $t+1$. Equation (3.41) can be written as

$$\begin{aligned} g^Y_t =& \frac{1}{Y_t} \sum_{i=1}^{N} \int_{I_{i,t}} \left[\Pi_{i,t+1} y_{t+1}(s,i)\phi_{t+1}(s|i)\mathrm{d}s - \Pi_{i,t} y_t(s,i)\phi_t(s|i)\mathrm{d}s \right] \\ & + \frac{1}{Y_t} \sum_{i=1}^{N} \int_{E_{i,t}} \left[\Pi_{i,t+1} y_{t+1}(s,i)\phi_{t+1}(s|i)\mathrm{d}s - \Pi_{i,t} y_t(s,i)\phi_t(s|i)\mathrm{d}s \right] \\ & + \frac{1}{Y_t} \sum_{i=1}^{N} \int_{G_{i,t}} \left[\Pi_{i,t+1} y_{t+1}(s,i)\phi_{t+1}(s|i)\mathrm{d}s - \Pi_{i,t} y_t(s,i)\phi_t(s|i)\mathrm{d}s \right], \end{aligned} \tag{3.42}$$

where the first term captures the GDP growth through the intensive margin, due to entrepreneurs within the credit regime expanding their businesses; the second term captures the GDP growth through the extensive margin, due to entrepreneurs' transition between the credit and non-credit regimes; and the last term reflects the general equilibrium effect.

Growth Accounting by Markets. The sets Ψ_0 and Ψ_0^C includes all markets with and without bank branches in 1986 (at $t = 0$), respectively. The economy's GDP at time t can be decomposed into

$$Y_t = \sum_{i \in \Psi_0} \Pi_{i,t} \int y_t(s,i)\phi_t(s|i)\mathrm{d}s + \sum_{i \in \Psi_0^C} \Pi_{i,t} \int y_t(s,i)\phi_t(s|i)\mathrm{d}s. \tag{3.43}$$



Note: Panel A decomposes the GDP growth by factor (labor, capital, and TFP) according to a standard growth accounting exercise. Panel B decomposes the GDP growth by credit regime as in Dabla-Norris et al. (2021). The extensive margin refers to the GDP growth resulting from entrepreneurs obtaining credit access after bank expansion. The intensive margin refers to the GDP growth resulting from entrepreneurs who already had credit access expanding their production after bank expansion. The general equilibrium effect (GE) is a residual term. Panel C decomposes the GDP growth by market, i.e., the markets without branches in 1986 versus those with branches in 1986. The formulas for the decomposition in each panel are provided in Online Appendix 3.6.1.

Figure OA.10: Decomposition of GDP growth by factor, credit regime, and market.

Thus,

$$\begin{aligned} g_t^Y =& \frac{1}{Y_t} \sum_{i \in \Psi_0} \left[\Pi_{i,t+1} \int y_{t+1}(s,i) \phi_{t+1}(s|i) \mathrm{d}s - \Pi_{i,t} \int y_t(s,i) \phi_t(s|i) \mathrm{d}s \right] \\ &+ \frac{1}{Y_t} \sum_{i \in \Psi_0^C} \left[\Pi_{i,t+1} \int y_{t+1}(s,i) \phi_{t+1}(s|i) \mathrm{d}s - \Pi_{i,t} \int y_t(s,i) \phi_t(s|i) \mathrm{d}s \right], \end{aligned} \tag{3.44}$$

where the first term captures the GDP growth attributed to the markets that have a branch in 1986, and the second term captures the GDP growth attributed to the markets that do not have a branch in 1986.

3.6.2 Results

Panel A of Figure OA.10 decomposes the GDP growth from 1986 to 1996 by factor (the formulas are presented in Online Appendix 3.6.1). A major part (52.4%) of the 26.5% GDP growth implied by the model is attributable to labor growth, which coincides with the pattern in panel C of Figure 8. About 26.0% of GDP growth is attributable to capital growth, and the remainder is attributable to TFP growth.

Panel B of Figure OA.10 decomposes the GDP growth by credit regime. Bank expansion promotes GDP growth by enabling entrepreneurs who were previously excluded from the credit market to borrow loans (extensive margin) and entrepreneurs who already had credit access to expand their businesses (intensive margin). The former contributes about 67.7% of GDP growth

and the latter contributes about 21.1%. The more important role played by the extensive margin is consistent with the dramatic increase in the credit access ratio in both the model and the data (panel D of Figure 8).

In panel C of Figure OA.10, we further decompose GDP growth according to geographical areas. Specifically, we separately consider the contribution to GDP growth of markets with branches in 1986 and that of markets without branches in 1986. We find that the latter group of markets accounts for almost all of the GDP growth, indicating the important role played by new branch openings. The former group experiences a slight positive growth, 2.1%, due to the migration inflows of talented entrepreneurs.

3.7 Discussions for the Prediction of Branch Locations

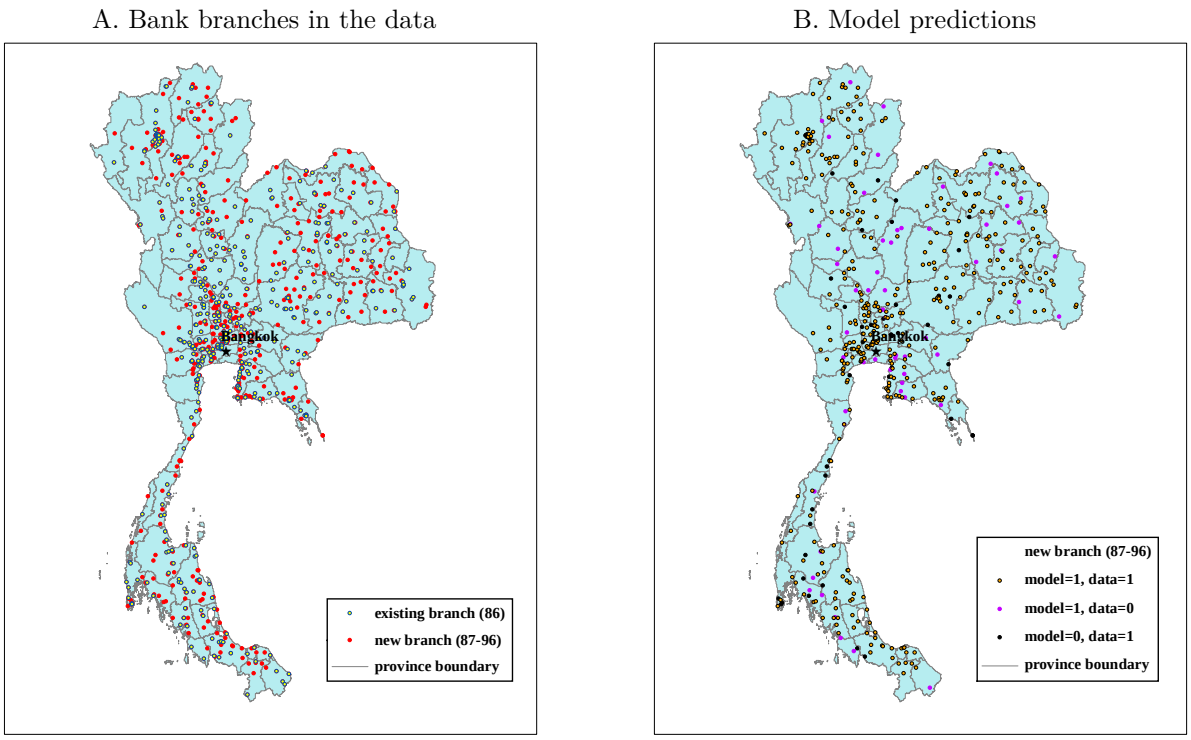
In this section of the appendix, we discuss our model's predictions of branch locations in more detail. Subsection 3.7.1 evaluates the model's prediction on branch locations against the data in 1996. Subsection 3.7.2 discusses the role of market characteristics in determining new branch locations. Subsection 3.7.3 discusses the implications of spatial spillovers for new branch locations. Subsection 3.7.4 discusses the spatial diffusion patterns of branches in the data and model. Subsection 3.7.5 evaluates the model's fit of branch locations in several major regions of Thailand. Subsection 3.7.6 further evaluates the model's fit of branch locations at the province level.

3.7.1 Locations of New Branch Openings

We compare the model-predicted branch locations in 1996 with those in the data in Figure OA.11. Panel A displays the branch locations in the data. The yellow dots represent the 406 branch locations in 1986, which are taken as given in our model calibration. The red dots represent the locations of 431 new branches opened from 1987 to 1996. Panel B assesses the predictive power of the model. Our model correctly predicts the locations of 372 (gold dots) of the 431 new branches opened from 1987 to 1996, with a correct prediction ratio of 86.3% (= 372/431).

3.7.2 Role of Market Heterogeneity

In this subsection, we focus on discussing the role of each of the market characteristics, $\Pi_{i,t}$, Z_i, and $d_{i,t}$, in determining branch locations in our model. As specified in equations (3.14) and (3.17) of the main text, the central authority chooses the locations of new branches to maximize the discounted profits from lending to households in the future, and these profits increase with households' demand for loans. Thus, to understand the predicted branch locations in our model, it is informative to examine how households' demand for loans varies in markets with different characteristics.

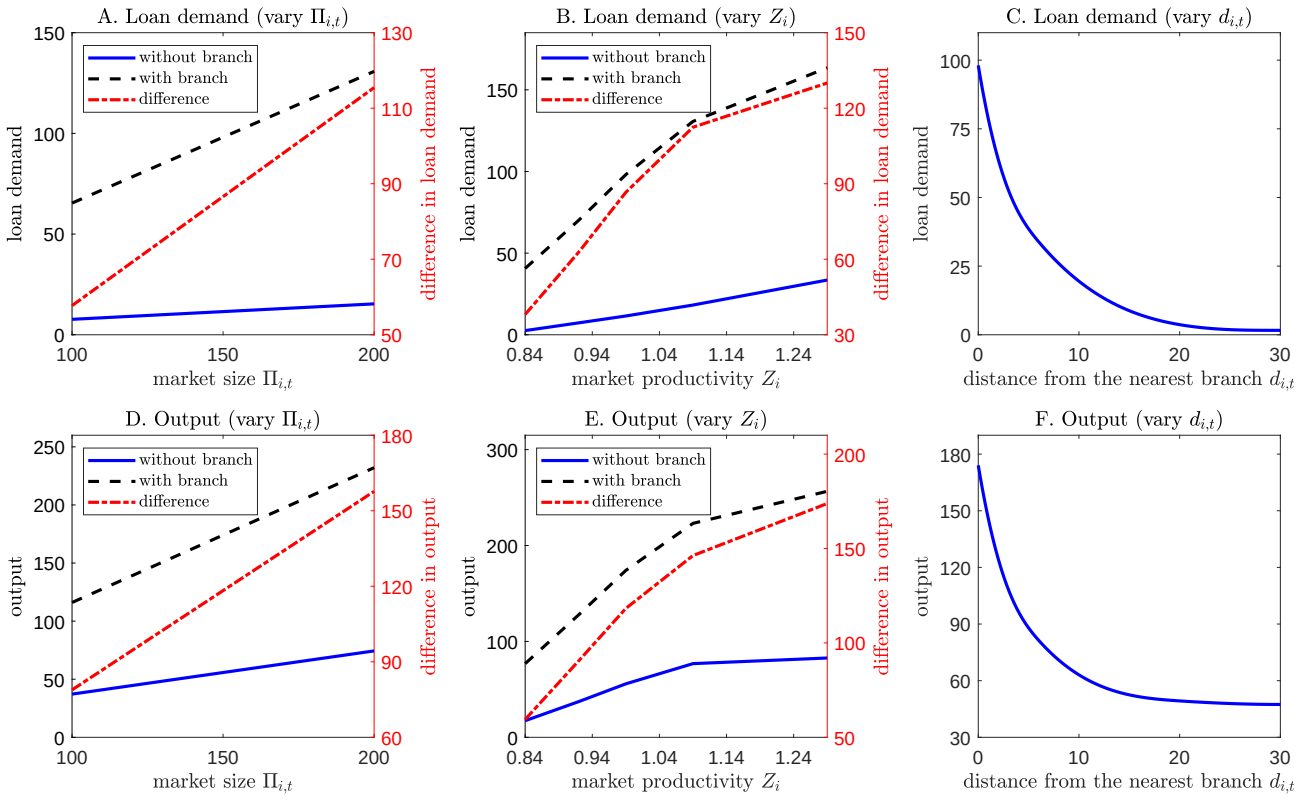


Note: Panel A displays the branch locations in the data. The yellow dots represent the 406 branch locations in 1986, which are taken as given in our model calibration. The red dots represent the locations of 431 new branches opened from 1987 to 1996. Panel B displays the model-predicted locations of the new branches opened from 1987 to 1996. The gold dots (372 in total) represent the branch locations for which the model's predictions are consistent with the data. The purple dots (59 dots) represent the locations where there are branches in the model but not in the data. The black dots (59 dots) represent the locations where there are branches in the data but not in the model.

Figure OA.11: Model vs. data: The distribution of bank branches.

Panel A of Figure OA.12 compares the per-period market-level loan demand in the steady state between a market without branches (blue solid line) and a market with a branch (black dashed line). Along the x-axis, we vary market size $\Pi_{i,t}$, holding productivity Z_i fixed at the median value across markets. Clearly, the market with a branch has a much higher loan demand than the market without branches owing to the lower credit entry costs $\psi_{i,t}$. As market size $\Pi_{i,t}$ increases, the loan demand in both markets increases simply because of a scaling effect. The difference between the two curves (the red dash-dotted line corresponding to the right y-axis) captures the increase in market-level loan demand when a branch is opened in the market. Importantly, it becomes larger as market size $\Pi_{i,t}$ increases. Because the central authority chooses branch locations to maximize total loan demand, panel A of Figure OA.12 suggests that, all else being equal, the central authority chooses to open branches in markets where $\Pi_{i,t}$ is larger.

Panel B of Figure OA.12 compares the two markets' loan demands for different productivity



Note: Panels A and B plot market-level demand for loans in a market without branches (blue solid line, $d_{i,t} = 12.8$ minutes, corresponding to the median car travel time to the nearest branch across all markets without branches in 1986) and a market with a branch (black dashed line) as a function of market size $\Pi_{i,t}$ and productivity Z_i, respectively. The red dash-dotted line plots the difference in loan demand between the two markets (corresponding to the right y-axis). Panel C plots the market-level demand for loans as a function of its distance from the nearest bank branch $d_{i,t}$. Panels D to F plot the market-level output corresponding to the markets in panels A to C, respectively. In panels A and D, we set the two markets' productivity at $Z_i = 0.99$, corresponding to the median productivity across all markets. In panels B and E, we set the two markets' size at $\Pi_{i,t} = 150$, corresponding to the median value across all markets. In panels C and F, we set market size and productivity at their median values. All parameter values are set according to our calibration in Table 2 of the main text.

Figure OA.12: Implications of $\Pi_{i,t}$, Z_t, and $d_{i,t}$ for a market's loan demand and output.

levels Z_i, holding market size $\Pi_{i,t}$ at the median value across markets. Similar to panel A of Figure OA.12, the market with a branch has a much higher loan demand than the market without branches. As productivity Z_i increases, the loan demand in both markets increases because a higher Z_i increases the return to entrepreneurial businesses, motivating more households to become entrepreneurs and to borrow. The red dash-dotted line (corresponding to the right y-axis) plots the difference between the two markets, which captures the increase in market-level loan demand when a branch is opened in the market. The line is upward sloping, indicating that this demand increases more when branches are opened in markets with higher productivity Z_i. This is because in markets with higher productivity Z_i, entrepreneurs prefer to expand and run their

businesses on a larger scale. Thus, they demand more bank loans once they obtain access to credit. All else being equal, panel B of Figure OA.12 suggests that the central authority chooses branch locations in markets with higher productivity Z_i.

Panel C of Figure OA.12 plots the market-level demand for loans as a function of the distance from the nearest branch $d_{i,t}$, holding market size $\Pi_{i,t}$ and productivity Z_i at their median values. The demand for loans strictly decreases as $d_{i,t}$ increases because of the higher credit entry costs $\psi_{i,t}$. This suggests that all else being equal, the central authority will choose to open branches in markets with a longer distance from the nearest branch $d_{i,t}$, as the increase in market-level loan demand in response to a branch opening is the highest in these markets compared with markets closer to bank branches.

In summary, panels A to C of Figure OA.12 suggest that in our model, because the central authority essentially chooses branch locations to maximize total loan demand, it prefers to open branches in markets with a larger size $\Pi_{i,t}$, that have higher productivity Z_i, or are farther away from existing branches (i.e., higher $d_{i,t}$). Once calibrated against the data, these three market characteristics are correlated with each other and jointly determine the choice of branch locations. In addition, because a newly opened branch can serve households not just in the market in which the branch is established, but also households living in nearby markets without branches (see the second term in the bracket of equation (3.17) in the main text), the central authority will take into account such spillover effects and consider the whole branch network in the optimization problem. We discuss the role of the branch network in Online Appendix 3.7.3 below.

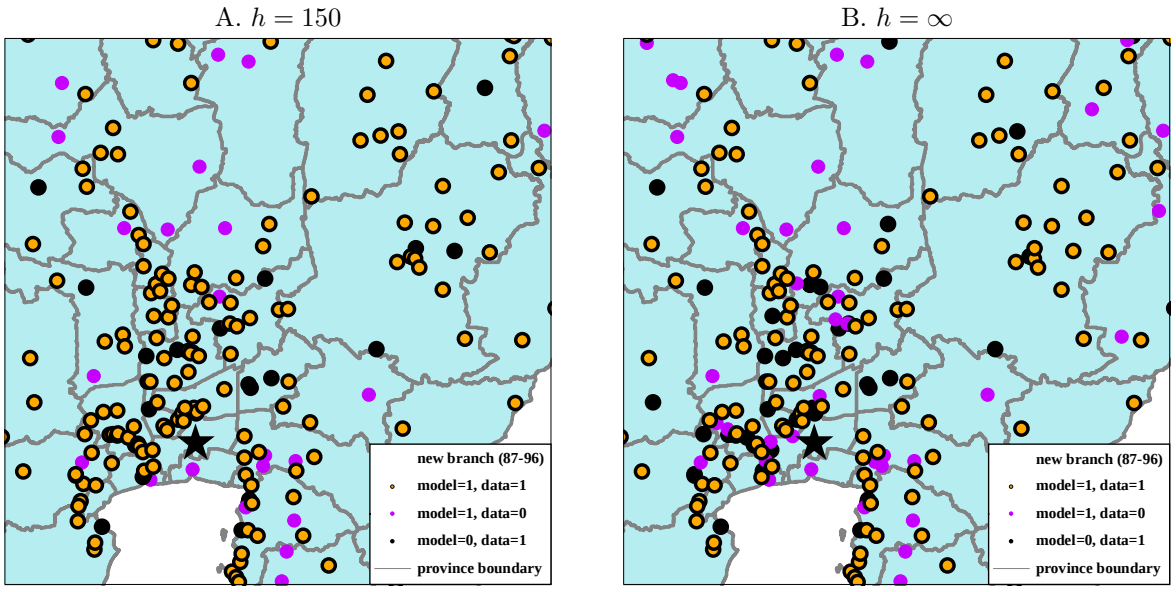
Panels D to F of Figure OA.12 plot the market-level output corresponding to the markets illustrated in panels A to C. The patterns of output are similar to those of loan demand: all else being equal, markets of larger size $\Pi_{i,t}$, with higher productivity Z_i, and located a shorter distance from the nearest branch $d_{i,t}$ have higher output. Thus, by maximizing total loan demand, the central authority's choice of branch locations also leads to significant GDP growth.

3.7.3 Spatial Spillovers and the Role of Branch Network

In our model, as noted above, opening a branch can serve not only households in the particular market in which the branch is located but also households living in nearby markets without branches (see the second term in the bracket of equation (3.17) of the main text). Therefore, the central authority will take into account such spatial spillovers and consider the whole branch network in the optimization problem. The degree of spatial spillovers is controlled by the branch capacity parameter h.[15]

Specifically, setting the branch capacity at $h = 0$ would restrict households from traveling to

[15]Given h, the degree of spatial spillovers is determined by the parameters θ_ψ and θ_ζ, governing the sensitivity of credit entry costs $\psi_{i,t}$ and portfolio adjustment costs $\zeta_{i,t}$ to the distance from the nearest bank branch $d_{i,t}$.



Note: This figure illustrates the impact of spatial spillovers on model-predicted bank branch locations near Bangkok. In panel A, we set the capacity $h = 150$ as in our calibration in Table 2 of the main text, meaning that households can travel to nearby markets to obtain bank services, but the size of these markets must be smaller than 150 people per square kilometer. In panel B, we set the capacity $h = \infty$, meaning that households can freely travel to nearby markets to obtain bank services. The spatial spillovers are the strongest in this case. The gold dots represent the locations of branches opened from 1987 to 1996 for which our model's predictions are consistent with the data. The purple dots represent the locations where there are branches in the model, but not in the data. The black dots represent the locations where there are branches in the data that are not predicted by the model.

Figure OA.13: Model-predicted bank branch locations near Bangkok.

other markets to obtain bank services (see equation (3.1) in the main text). Thus, setting $h = 0$ is equivalent to setting $\tau_{ij} = \infty$ for all $i \neq j$, which we present as a counterfactual experiment in Table 3 of the main text (see column (2)). In another extreme case, when $h = \infty$, bank branches do not have any capacity constraints in terms of serving customers. Thus, households may travel to any other market to obtain loans, without restrictions, maximizing the impact of spatial spillovers.[16] In an intermediate case with $h \in (0, \infty)$, households can travel to markets whose market size $\Pi_{i,t}$ is less than h, enabling the model to capture partial spatial spillovers. Focusing on a specific region of interest, such spatial spillovers will lead to fewer branches being opened in the region (given that households can use nearby branches). Thus, whether increasing branch capacity h improves or reduces the model's predication accuracy for a region depends on whether the model predicts more or fewer branches relative to the data.[17]

To elaborate on the spatial spillover within a cluster and its effect on our prediction, we focus

[16] These households of course need to pay the credit entry cost $\psi_{i,t}$ and portfolio adjustment cost $\zeta_{i,t}$, both of which increase with their travel distance (see our calibration in Table 2 of the main text).

[17] In principle, it is possible to further improve the prediction of branch locations in each region by incorporating more sophisticated branch capacity constraints, such as by allowing h to differ across markets.

on the model predictions for the Bangkok metropolitan area in Figure OA.13. Panel A presents the model-predicted branches when $h = 150$, as in our calibration in Table 2 of the main text. Panel B presents the model-predicted branches when $h = \infty$, with no branch capacity constraint (i.e., no restriction on households' traveling to other markets to obtain bank services). Comparing the two panels, it is clear that adding the branch capacity constraint leads the model to predict that more branches will be located in the cluster of markets around Bangkok than elsewhere, which is consistent with the data (i.e., there are more gold dots around Bangkok in panel A than in panel B).

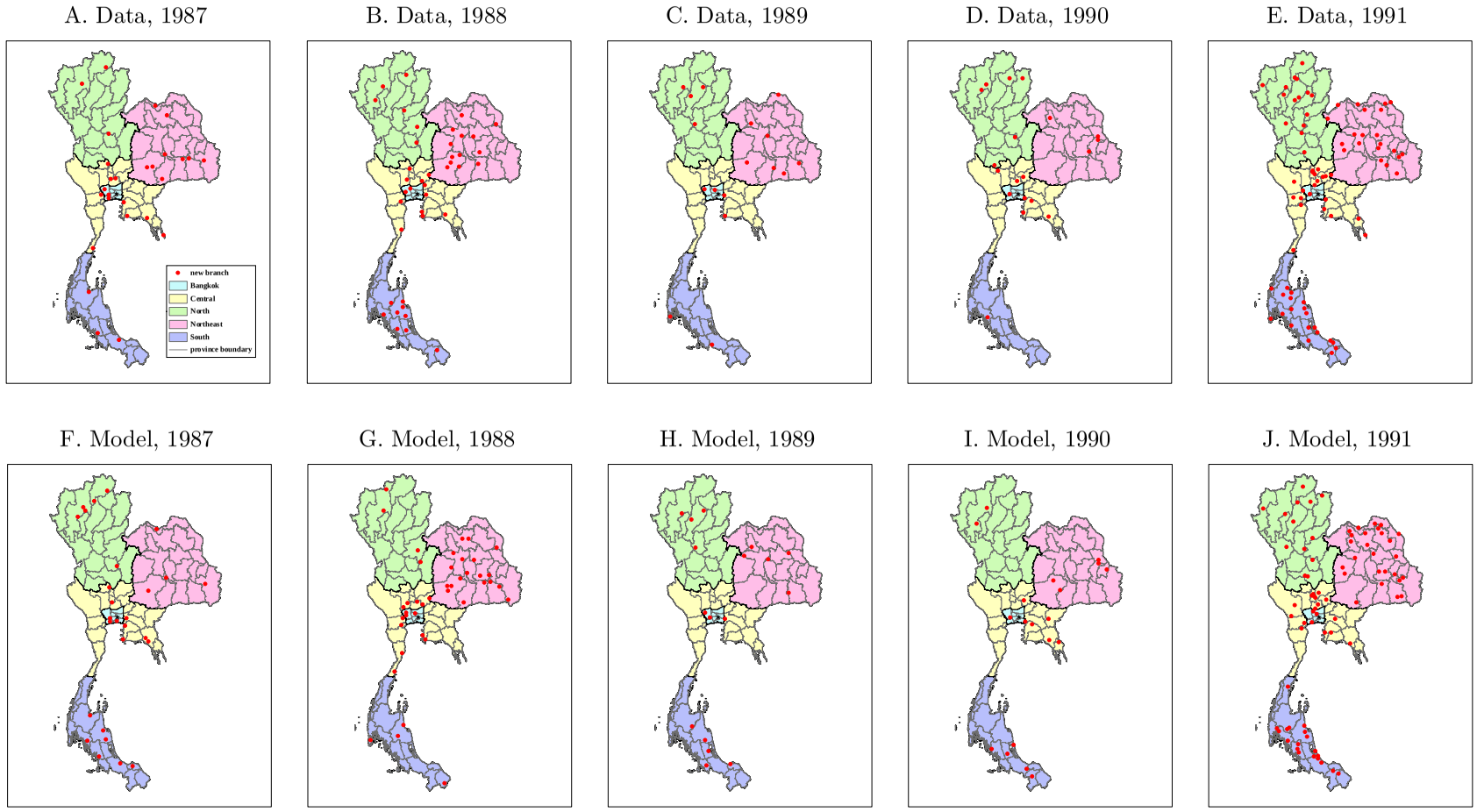
3.7.4 Spatial Diffusion Patterns of Branches

In this subsection, we examine whether the spatial diffusion pattern of bank branches is similar to that of retail businesses. In a seminal work, Holmes (2011) studies the expansion of Wal-Mart and finds that it always places new stores close to each other. That is, Wal-Mart always opens stores close to existing stores, and never locates stores in far-off locations with the intent of later filling in the area in between.

Figures OA.14 and OA.15 display the bank branches opened in each year during 1987-1996. At first glance, it is difficult to see a spatial diffusion pattern that resembles the expansion of Wal-Mart documented by Holmes (2011). The new branches opened each year are located far from each other and spread over the entire country. Many new branches opened in the same year are located in different provinces. In contrast to the spatial diffusion pattern displayed in Figures 1 and 2 of Holmes (2011) for Wal-Mart, the expansion of bank branches in Thailand does not appear to begin at a certain point and radiate from the center outward.

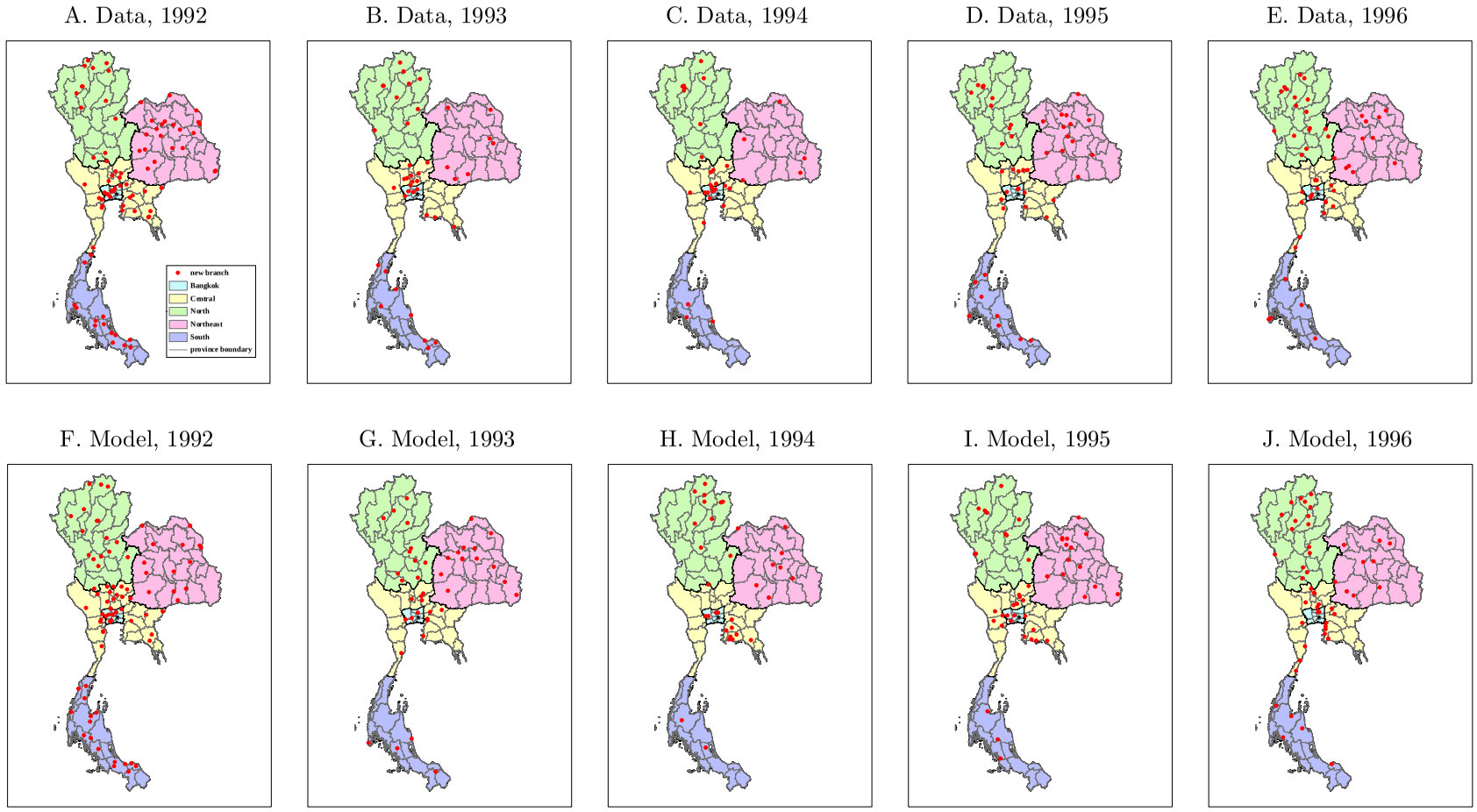
To more clearly investigate whether bank branches are opened close to existing ones, we construct a geographic proximity measure for the closeness of newly opened branches to existing bank branches. In particular, for each year $t = 1987, ..., 1996$, we calculate the average distance of n_t new branches opened in year t from the nearest existing branches opened before year t. The blue solid line in panel A of Figure OA.16 shows that for the 1987-1991 period, the average distance between a new branch and the nearest existing branch is about 30-40 minutes (in terms of car travel time). Toward the end of our study period, during 1992-1996, this average distance decreases to 16-25 minutes, given that many more markets have branches in these years compared with earlier years. As a benchmark, the black dashed line in panel A of Figure OA.16 plots the median value of the distribution of the distance from the nearest branch for all markets without branches in each year. Comparing the blue solid and black dashed lines in each year t, the average distance of new branches from the nearest existing branches exceeds the median value of the distribution, being 2 to 3 times the median value.

Next, we construct the distribution of the distance from the nearest branch across all the



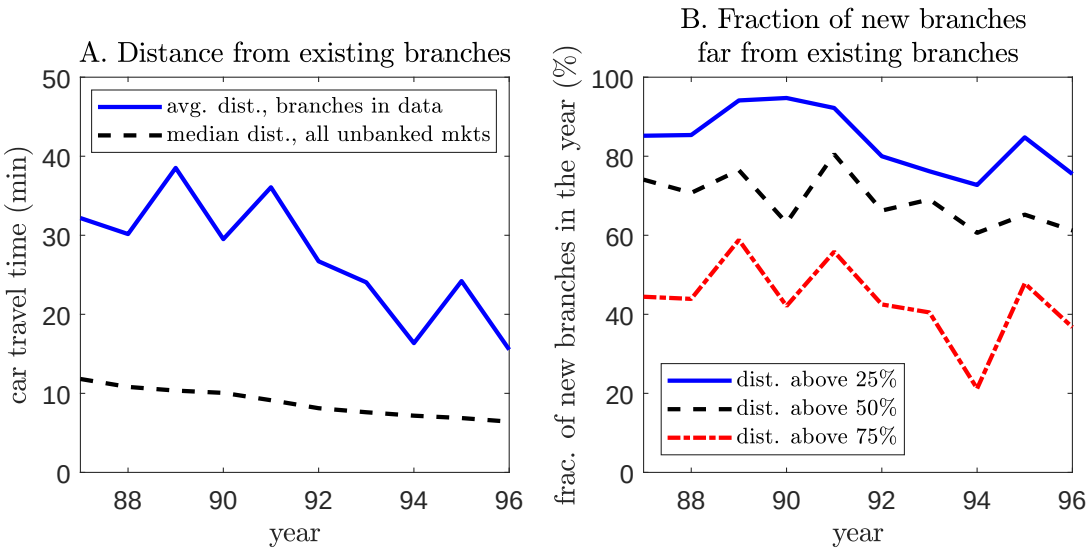
Note: This figure displays the new branches opened in each year from 1987 to 1991. Panels A to E refer to the data and panels F to J refer to the model.

Figure OA.14: Branches opened in each year between 1987 and 1991.



Note: This figure displays the new branches opened in each year from 1992 to 1996. Panels A to E refer to the data and panels F to J refer to the model.

Figure OA.15: Branches opened in each year between 1992 and 1996.



Note: In panel A, the blue solid line plots the average distance of n_t new branches opened in year $t = 1987, ..., 1996$ from the nearest existing branches opened before year t. The black dashed line plots the median value of the distribution of the distance from the nearest branch across all markets without branches in each year. In panel B, the blue solid, black dashed, and red dash-dotted lines plot the fraction of new branches whose distance from the nearest existing branches is above the 25%, 50%, and 75% points, respectively, of the distribution of the distance from the nearest branch across all the markets without branches in each year.

Figure OA.16: Distance of new branches opened in each year from existing branches.

markets without branches in each year. Panel B of Figure OA.16 plots the proportion of new branches whose distance from the nearest existing branches is above the 25th (blue solid line), 50th (black dashed line), and 75th (red dash-dotted line) percentile points in the distribution in each year. The figure shows that about 70% of the new branch locations are in the top 50% of the distribution (blue solid line). More than 40% of the branches are opened in far-off locations, indicated by their distance from the nearest branch being above 75% in the distribution (red dash-dotted line). This analysis further confirms that, indeed, many new branches are opened in locations that are distant from existing branches in Thailand during this period.

In summary, we find that the spatial diffusion pattern of bank branches in our data is very different from that of retail businesses, such as Wal-Mart, which benefit from economies of density. The essential reason for Wal-Mart's strategy, as suggested by Holmes (2011), is that it gains logistic benefits by operating its stores close to each other. Products sold by Wal-Mart are supplied by regional distribution centers. When stores are opened close to a distribution center, Wal-Mart can save on transport costs and quickly respond to demand shocks. However, in the case of bank branches, such logistical issues are not first-order concerns because both bank deposits and loans are managed through an electronic funds transfer system. The lack of economies of density appears to justify why bank branches are opened in regions remote from existing branches, which ensures that their market areas do not overlap significantly.

Rysman, Townsend and Walsh (2022) conduct formal empirical tests by estimating an ordered probit regression model that distinguishes the top banks in Thailand by name. Their regression results provide some empirical evidence that suggests a similar spatial pattern to the one we find. Banks are less likely than retail stores to open new branches in a market that already has branches. Similar patterns are also documented for the U.S. Following the interstate banking deregulation, banks choose to open branches in other states, far away from their existing branches. Goetz, Laeven and Levine (2016) provide evidence that such out-of-state geographic expansion reduces banks' exposure to idiosyncratic local risks.

3.7.5 Predictions at the Regional Level

To assess the geographic patterns of branch expansion for the 1987-1996 period, we divide Thailand into five regions, comprising the Bangkok metropolitan, central, north, northeast, and south regions, based on the 1996 SES in Thailand.

Table OA.11 summaries the number of branches opened in each region and each year, as displayed in Figures OA.14 and OA.15. The numbers in the data and baseline model are visualized in Figure 6 of the main text. In addition, the table presents the prediction of an alternative model specification, under which all markets have the same level of productivity (achieved by setting $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$). Compared with the data, this alternative model without market-level productivity heterogeneity can still predict the number of branches opened in the north and south well. However, it predicts too many branches in the northeast and too few branches in the Bangkok metropolitan and central regions. Intuitively, the markets in the Bangkok and central regions are estimated to have higher productivity than markets in other regions (see panel B of Figure OA.3 in Online Appendix 1.2). By contrast, markets in the northeast, although populous, are not very productive. As we discuss in Online Appendix 3.7.2, in our model, branches tend to be opened in markets with higher productivity Z_i because of the goal of maximizing profits. Thus, if the model does not feature heterogeneity in market productivity, it will predict too few branches around Bangkok but too many in the northeast.

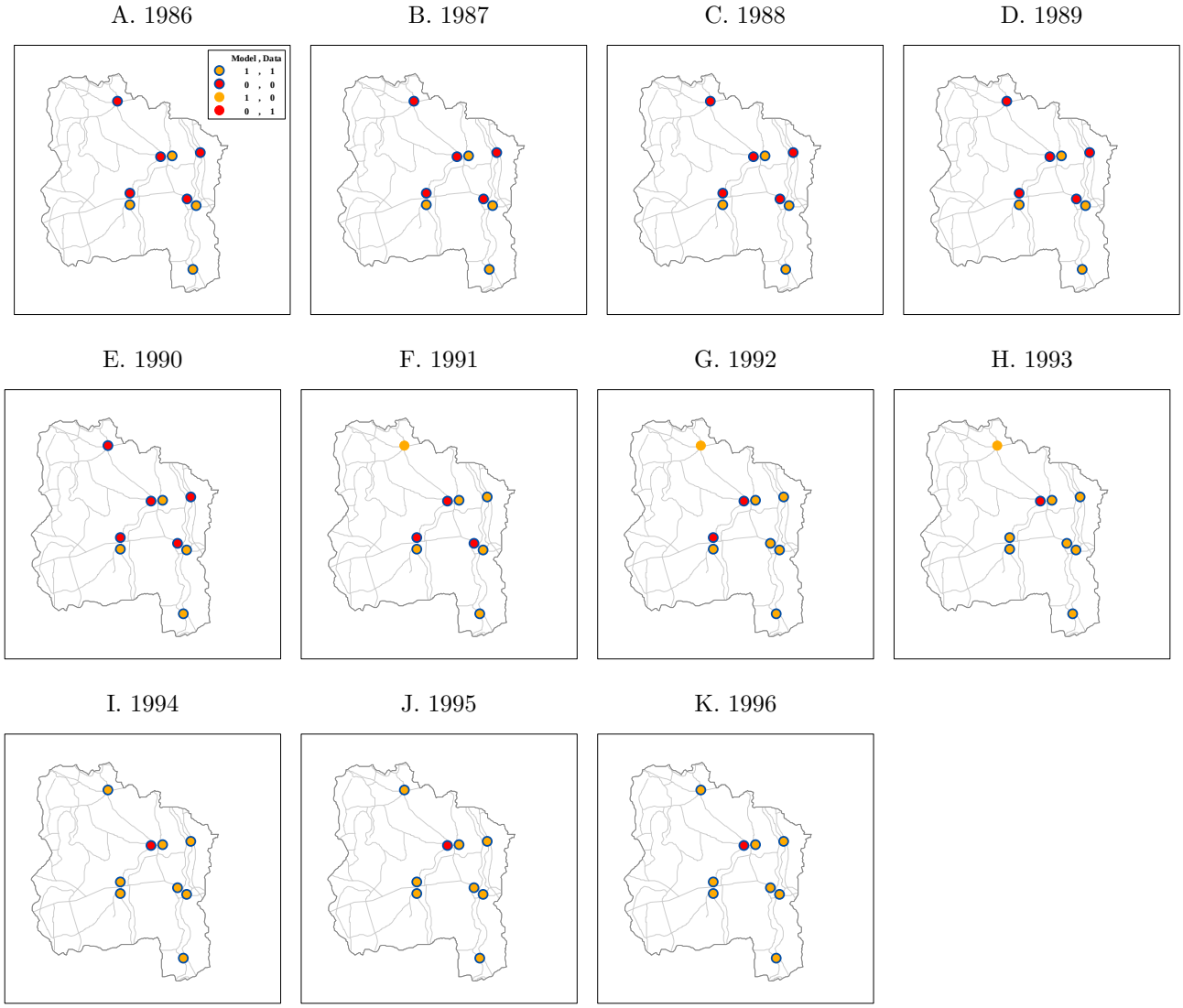
3.7.6 Predictions at the Province Level

Figures OA.17 and OA.18 report the dynamics of bank expansion in the two provinces corresponding to the 25th and 75th percentiles (Ang Thong and Sa Kaeo, respectively) of the distribution of province-level average timing differences (see Figure 7 of the main text), respectively.

Table OA.11: Branches opened in different regions between 1987 and 1996.

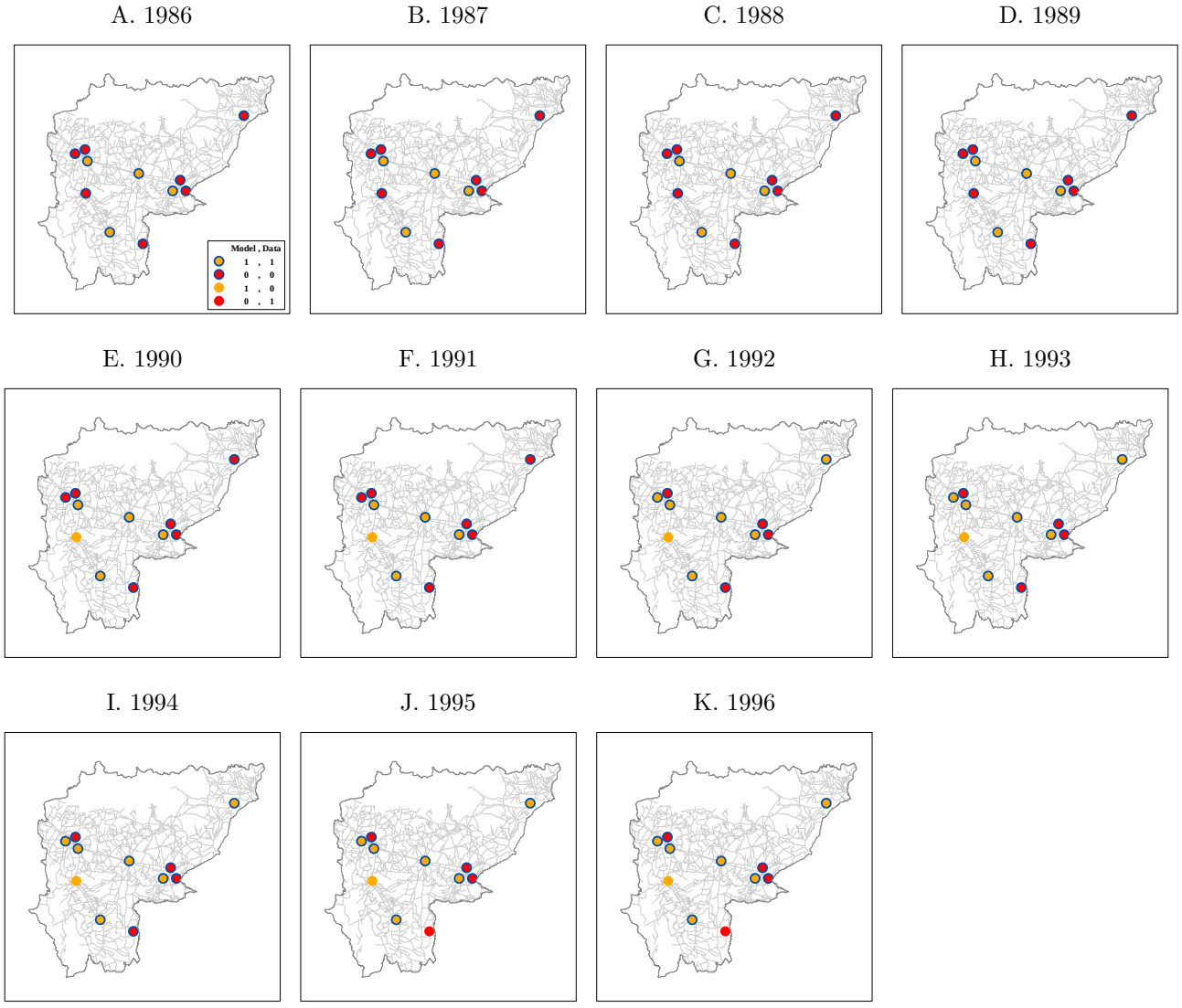
		1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	Cumulative
Bangkok metropolitan												
Data		3	2	2	1	1	8	3	6	3	3	32
Model	baseline	3	3	2	1	3	7	2	4	3	2	30
	$Z_i \equiv \overline{Z}$	2	0	0	1	5	4	1	1	4	5	23
Central												
Data		9	11	2	7	19	23	14	10	11	13	119
Model	baseline	7	11	1	6	16	24	11	9	16	17	118
	$Z_i \equiv \overline{Z}$	3	3	3	2	19	22	11	5	11	7	86
North												
Data		3	6	4	6	15	13	10	9	11	17	94
Model	baseline	6	4	4	2	13	18	11	10	10	16	94
	$Z_i \equiv \overline{Z}$	2	7	3	6	15	19	13	14	7	10	96
Northeast												
Data		9	13	7	4	22	22	7	5	13	11	113
Model	baseline	4	19	5	5	26	15	14	8	14	9	119
	$Z_i \equiv \overline{Z}$	19	27	10	4	25	22	8	3	20	22	160
South												
Data		3	9	2	1	20	14	8	3	8	5	73
Model	baseline	7	4	5	5	19	16	4	2	3	5	70
	$Z_i \equiv \overline{Z}$	1	4	1	6	13	13	9	10	4	5	66
Total		27	41	17	19	77	80	42	33	46	49	431

Note: Based on Figures OA.14 and OA.15, we count the number of branches opened in each region and each year. The rows labeled by "baseline" present our baseline model's prediction. The rows labeled by "$Z_i \equiv \overline{Z}$" present the prediction of an alternative model specification in which all markets have the same productivity (by setting $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$).



Note: This figure compares the model-predicted and actual bank branch locations from 1986 to 1996 in Ang Thong province, which corresponds to the 25th percentile of the distribution of province-level timing differences. Each circle represents the center of a market. A circle has a blue edge and is filled with gold if there is a branch in both the model and data $(1, 1)$. A circle has a blue edge and is filled with red if there is no branch in both the model and data $(0, 0)$. A circle without an edge but is filled with gold indicates that there is a branch predicted by the model but not in the data $(1, 0)$. Finally, a circle without an edge but is filled with red indicates that there is a branch in the data but not predicted by the model $(0, 1)$. The average timing difference between the model-predicted and actual bank locations is 0.75 years in this province. The thin gray lines in each panel represent the road network.

Figure OA.17: Model-predicted and actual bank branch locations in Ang Thong.



Note: This figure compares the model-predicted and actual bank branch locations from 1986 to 1996 in Sa Kaeo province, which corresponds to the 75th percentile of the distribution of province-level timing differences. Each circle represents the center of a market. A circle has a blue edge and is filled with gold if there is a branch in both the model and data $(1, 1)$. A circle has a blue edge and is filled with red if there is no branch in both the model and data $(0, 0)$. A circle without an edge but is filled with gold indicates that there is a branch predicted by the model but not in the data $(1, 0)$. Finally, a circle without an edge but is filled with red indicates that there is a branch in the data but not predicted by the model $(0, 1)$. The average timing difference between the model-predicted and actual bank locations is 3.33 years in this province. The thin gray lines in each panel represent the road network. The white area in the northern region of the province is occupied by a large mountain, so the road network does not cross this area.

Figure OA.18: Model-predicted and actual bank branch locations in Sa Kaeo.

3.8 Discussions on the Choice of Branch Locations

In this section of the appendix, we discuss the choice of branch locations in more depth. In Subsection 3.8.1, we formulate the choice of branch locations in an alternative way to the main text to capture a banking sector that is more competitive than a national monopoly. In Subsection 3.8.2, we compare the predictions of branch locations in a model with this alternatively formulated banking sector with those of our baseline model with a monopoly banking sector. In Subsection 3.8.3, we discuss the fact that the two models have similar implications because the spatial spillovers of branches on nearby markets (or cannibalization across branches) are not an important driver for branch locations.

3.8.1 An Alternative Formulation for the Choice of Branch Locations

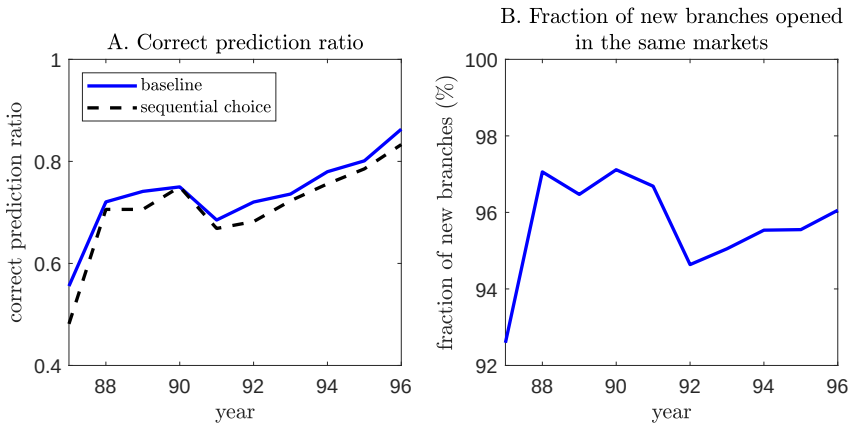
In our baseline model (Section 3.3 of the main text), the locations of all new branches, $\{\Lambda_t\}_{t=1}^{\infty}$, are determined at the very beginning ($t = 0$) by a central authority to maximize total profits (equation (3.14) of the main text) as if there is a coordinating profit-maximizing monopoly. Here, we consider an alternative formulation for the choice of branch locations. Specifically, we assume that at the beginning of each period $t \geq 1$, the central authority only chooses the locations of n_t branches opened in the current period to maximize the total profits of these n_t branches, as follows:

$$\max_{\Lambda_t} \quad \sum_{i \in \Lambda_t} \Theta_{i,t}, \tag{3.45}$$

$$\text{s.t.} \quad \Lambda_t \subset \Psi_{t-1}^{C} \quad \text{and} \quad |\Lambda_t| = n_t. \tag{3.46}$$

The set of markets with branches, Ψ_t, evolves as in equation (3.16) of the main text. A comparison of equation (3.45) with equation (3.14) of the main text indicates that the only difference is that, the central authority chooses Λ_t sequentially at the beginning of each period ($t \geq 1$) here, instead of all $\{\Lambda_t\}_{t=1}^{\infty}$ at the very beginning ($t = 0$) in the main text. However, we emphasize that equation (3.45) does not imply that the central authority is myopic because $\Theta_{i,t}$ represents the present value of future profits generated by the branch in market i in all future periods ($t' \geq t$) in the equilibrium with rational expectations, as defined in equation (3.17) of the main text.

Equation (3.45) can be interpreted in two ways. First, it can be regarded as capturing a central authority with limited commitment, which can only decide the locations of n_t branches to be opened in one period. Second, it may be seen as capturing a more competitive banking sector than a sector under one authority, in that the locations of branches opened in different periods are chosen by distinct central authorities. That is, the time-t central authority chooses the locations of n_t branches, rationally anticipating the locations of $\sum_{s=t+1}^{\infty} n_s$ branches opened



Note: In panel A, the blue solid line plots our baseline model's correct prediction ratio for new branches opened between 1987 and 1996. The black dashed line plots the correct prediction ratio of an alternative model using formulation (3.45) to determine new branch locations, i.e., the locations of branches opened in each period are sequentially chosen by a distinct central authority.

Figure OA.19: Comparing the baseline model with a model with an alternative formulation for the location choice of new branches opened between 1987 and 1996.

by future central authorities (i.e., time-s central authorities for all $s > t$), which can possibly reduce the profits of n_t branches opened in period t.

3.8.2 Comparison with the Baseline Model

We calibrate a model using equation (3.45) for the banking sector to match the same moments in Table 2 in the main text. Comparing this alternative model with our baseline model with a monopoly banking sector, we find that the two models produce very similar predictions of branch locations. Specifically, in panel A of Figure OA.19, the blue solid line plots our baseline model's correct prediction ratio for new branches opened between 1987 and 1996. The black dashed line plots the correct prediction ratio of this alternative model using equation (3.45), where the locations of branches opened in each period are sequentially chosen by distinct central authorities. The two lines are very close to each other, although our baseline model has a slightly higher correct prediction ratio than the alternative model. For example, in 1996, the correct prediction ratio is 86.3% for our baseline model and 83.3% for the alternative model.

Next, we evaluate the discrepancy in the predicted branch locations between the two models. Specifically, in each year $t = 1987, ..., 1996$, for those markets where new branches are opened between 1987 and t in our baseline model, we calculate the fraction of markets that also have new branches opened between 1987 and t in the alternative model. Panel B of Figure OA.19, which plots this ratio for the 1987-1996 period, shows that, on average, about 95% of the new branch locations predicted by our baseline model are also predicted by the alternative model. For example, in 1996, of the 431 new branch openings between 1987 and 1996, 414 branches (96.1%)

are predicted to open in the same markets by the two models.

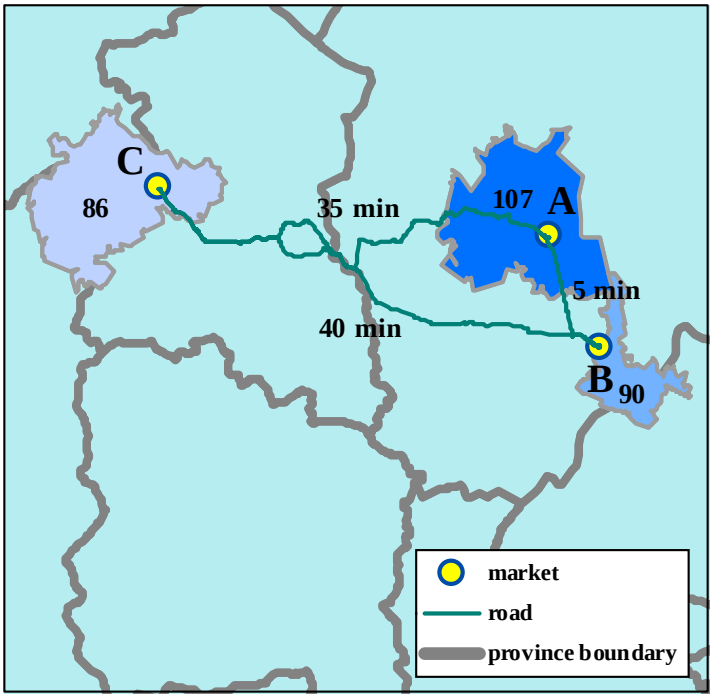
In terms of spatial patterns, we find that our baseline model tends to predict more branches in markets further away from existing branches than does the alternative model. As a result, our baseline model has a marginally better fit with the data than the alternative model. The main reason is that the central authority in our baseline model chooses all branch locations at the very beginning ($t = 0$) to maximize the total profits of all branches opened over the 1987-1996 period, rather than sequentially choosing branch locations in each period t only for those branches opened in the current period. To elaborate, Figure OA.20 provides one such example in our calibrated branch network. In the baseline model, the central authority chooses to open a branch in market A in 1991 and a branch in market C in 1992. The central authority does not open a branch in market B even though this market itself generates more profits than market C. This is because market B is close to market A, so households in market B can already borrow from the branch in market A, albeit at a higher credit entry cost than they would incur in the case of a branch in market B owing to travel costs. In other words, the central authority understands that opening a branch in market B reduces market A's profits due to cannibalization and therefore chooses to open a branch in market C instead to maximize total profits. By contrast, in the alternative model, the central authority making decisions in 1991 chooses to open a branch in market A in 1991, but the central authority making decisions in 1992 (which is different from the central authority in the previous year) chooses to open a branch in market B in 1992 to maximize its own profits, ignoring the fact that households in market B can already borrow from the branch in market A. In addition, once a branch is opened in market B in 1992, households in market B will not travel to market A to borrow. This consequently reduces the profits made by the branch in market A from 1992 onward, as anticipated by the central authority that opens the branch in market A in 1991.

3.8.3 Discussions

The two models have similar predictions for branch locations because the spatial spillovers of branches on nearby markets (or cannibalization across branches) are not an important driver for branch locations.

Intuitively, it is the interaction among branches, or spatial spillovers, that leads the model-predicted locations to vary depending on whether the locations are chosen jointly or sequentially. If there are no spatial spillovers, the profits generated by branches opened in different markets are independent from each other.[18] That is, the profits of opening a branch in market i do not depend on whether there are branches in other markets because the branch's profits are

[18]Mathematically, equation (3.17) of the main text becomes $\Theta_{i,t} = \chi \sum_{t'=t}^{\infty} \beta^{t'} X_{i,t'}$ in the absence of spatial spillovers, meaning that the profits from opening a branch in market i depend only on the loans made to local entrepreneurs in market i.

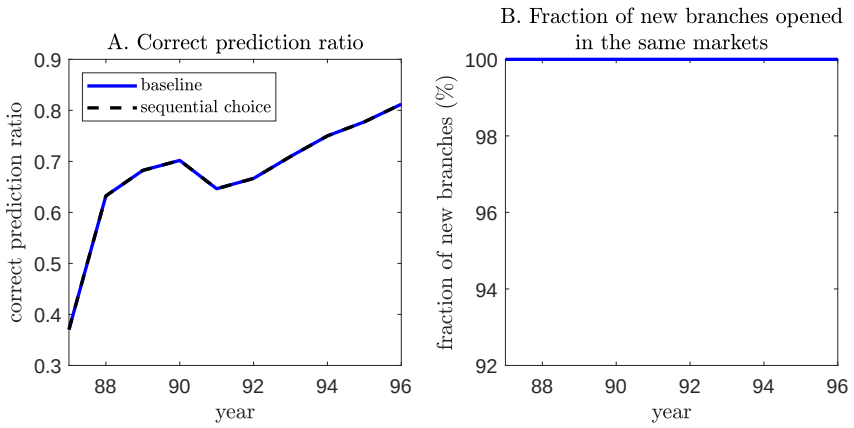


Note: There are three markets A, B, and C. The markets occupy a certain area in the data, as illustrated in panel A of Figure 2 in the main text. The yellow dots represent the locations of the markets, where branches can be opened. The car travel time is 5 minutes between markets A and B, 35 minutes between markets A and C, and 40 minutes between markets B and C. The number inside each market represents the discounted future profits that the market generates by itself, without considering spatial spillovers. That is, the profits of market i are computed according to $\Theta_{i,t} = \chi \sum_{t'=t}^{\infty} \beta^{t'} X_{i,t'}$ if a branch is opened in period t, where $X_{i,t'}$ is defined in equation (3.19) in the main text. The profits of markets A, B, and C are 107, 90, and 86, respectively.

Figure OA.20: An illustrative example for how spatial spillovers affect predicted branch locations.

completely generated by the loans made to entrepreneurs in local market i. Thus, in the absence of spatial spillovers, the central authority, regardless of whether it chooses all branch locations jointly or sequentially, should focus solely on opening branches in the markets that generate the highest profits by themselves individually.

Next, we explain why spatial spillovers are not an important driver for branch locations in our calibrated baseline model. In Table 3 of Section 5.2 of the main text, we show that setting $\tau_{ij} = \infty$, which essentially eliminates spatial spillovers in our model, only reduces the correct prediction ratio in 1996 from 86.3% (the baseline model with spatial spillovers) to 81.2%. This indicates that the role of spatial spillovers (which is related to the market characteristic $d_{i,t}$) is much less important than the two other market characteristics, size $\Pi_{i,t}$ and productivity Z_i, in explaining the branch locations. In both models, size $\Pi_{i,t}$ and productivity Z_i are calibrated similarly and these two market characteristics drive the locations of over 80% branches opened between 1987



Note: The plot is similar to Figure OA.19, except that we set $\theta_\psi = \infty$ in our calibration.

Figure OA.21: $\theta_\psi = \infty$: Comparing the baseline model with a model with an alternative formulations for the location choice of new branches opened between 1987 and 1996.

and 1996. Therefore, it is not surprising that the two models have similar predictions for branch locations, even though they adopt different formulations for the choice of branch locations.

Of course, the importance of spatial spillovers depends on the model calibration. Taking the network structure (i.e., τ_{ij} for all $i, j \in \mathcal{C}$) as given by the data, the key parameter that determines the importance of spatial spillovers is θ_ψ, which governs the sensitivity of the credit entry costs $\psi_{i,t}$ to the distance from the nearest branch $d_{i,t}$. The parameter θ_ψ directly determines entrepreneurs' demand for loans and thus the profits of bank branches. In our calibration, the parameter θ_ψ is calibrated to match the DID estimate for the impact of branch openings on credit access (see Table 2 in the main text). If we were to calibrate a higher θ_ψ, the spatial spillovers of bank branches on nearby markets would be less important, and, in this case, we would expect that the model-predicted branch locations would depend even less than under our baseline calibration on whether these locations are chosen jointly or sequentially.

As an example, consider an extreme calibration in which $\theta_\psi = \infty$, which essentially restricts households from borrowing from non-local branches. Equation (3.17) of the main text is simplified to $\Theta_{i,t} = \chi \sum_{t'=t}^{\infty} \beta^{t'} X_{i,t'}$, meaning that the profits from opening a branch in market i depend only on the loans made to local households in market i. Under this calibration, branches do not directly interact with each other,[19] and they have no spatial spillovers on nearby markets. Figure OA.21 shows that in this case, the two models with different formulations for the choice of branch locations yield identical predictions.

[19]Strictly speaking, branches still indirectly interact with each other through general equilibrium effects because opening a branch in one market can affect the other branches' profits due to the change in interest rates or wages. However, because there are 1,428 markets, the general equilibrium effect generated by changing one market's status is very small.

3.9 Anticipation and Lagged Effects of Branch Openings

When matching the treatment and control markets in our empirical analyses (Section 2.3.2 of the main text), we restrict the donor pool to include only markets that do not have any branch openings until $T = 1997$. Given that our estimation uses a sample for the 1986-1996 period, this restricted donor pool ensures that our estimates are not confounded by the lagged effect of past branch openings in control markets. Imposing such restrictions on the donor pool is standard in the empirical analyses with matching (e.g., Azoulay, Graff Zivin and Wang, 2010).

In this section of the appendix, we evaluate how the anticipation and lagged effects of branch openings could affect the DID estimates in the simulated data. Inspired by the experiments of Kaplan and Violante (2014), we evaluate how the anticipated and lagged effects of branch openings in control markets can potentially affect the DID estimates using the simulated data of our structural model. In this purely model-based exercise, we focus on the rational expectation equilibrium where all future branch openings are anticipated by households, consistent with the way in which we solve the model. In other words, we do not evaluate the implications of alternative information structures, which are partly the focus of Kaplan and Violante (2014).

Our main findings are summarized as follows. First, we find that the lagged effect of branch openings in control markets can significantly bias the DID estimates because it takes many years for control markets to reach the steady state after branch openings. However, we find that the anticipated effect of branch openings in control markets is not significant, except perhaps for a branch opened immediately in the next year. These model implications, if one is willing to take them seriously, justify our choice of donor pool in the empirical DID estimation based on the sample for the 1986-1996 period (Section 2.3.2 of the main text). That is, the donor pool is restricted to markets that do not have any branch openings until $T = 1997$, which completely rules out any potential lagged effect of past branch openings in control markets. Moreover, because our structural model suggests that the anticipated effect is not significant,[20] we do not further restrict the size of the donor pool by extending T to the years beyond 1997. We need a sufficiently large number of markets in our donor pool for matching purposes.

Table OA.12 reports the results. As a benchmark, column (1) of Table OA.12 reports the estimated α_0 (i.e., the effects in the first 2 years after opening a branch) in the simulated data. These estimates are slightly different from the estimates reported in Figure 5 of the main text because treatment and control markets are not matched by propensity scores. Instead, here, the DID coefficients are estimated based on pairs of treatment and control markets with perfectly identical characteristics. This allows us to more cleanly evaluate the anticipation and lagged effects of branch openings in the simulated data by comparing different columns in Table OA.12.

[20] In the data, the anticipated effect could be even weaker than in our model because branch openings may not be well anticipated by every household.

Table OA.12: Impacts of the anticipation and lagged effects of branch openings in control markets on DID estimates.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	benchmark	anticipation effects					lagged effects				
Branch opening time in control markets	N/A	$t+10$	$t+8$	$t+6$	$t+4$	$t+2$	$t-2$	$t-4$	$t-6$	$t-8$	$t-10$
ln(Employment)	0.212	0.212	0.212	0.212	0.210	0.207	0.110	0.151	0.192	0.201	0.205
ln(Income per capita)	0.259	0.259	0.259	0.259	0.256	0.252	0.145	0.190	0.234	0.246	0.250
Fraction of entrepreneurs	0.018	0.018	0.018	0.018	0.016	0.015	0.011	0.012	0.014	0.015	0.016
Fraction of HH with bank loans	0.032	0.032	0.032	0.032	0.032	0.032	0.026	0.032	0.032	0.032	0.032

Note: Column (1) reports the benchmark estimate for α_0 in the simulated data, where the DID estimation is based on pairs of treatment and control markets with perfectly identical characteristics. Specifically, in the simulated data, for each year $t = 1987, ..., 1996$, we first find all treatment markets (i.e., the markets in which a branch is opened in year t). Next, for each treatment market i in year t, we construct a synthetic control market i' with identical characteristics (i.e., $\Pi_{i,t-1} = \Pi_{i',t-1}$, $Z_i = Z_{i'}$, and $d_{i,t-1} = d_{i',t-1}$) in year $t-1$ prior to the treatment. We ensure that the synthetic control market neither had a branch opened before t nor will have a branch opened after t to ensure that our model-implied DID estimates are not confounded by the anticipation or lagged effects of branch openings in the synthetic control market. In columns (2) to (6), we evaluate how the anticipation effect of branch openings in control markets affect the DID estimates in the simulated data. Similar to what we do in column (1), each treatment market i in year t is matched with a synthetic control market i' of identical characteristics (i.e., $\Pi_{i,t-1} = \Pi_{i',t-1}$, $Z_i = Z_{i'}$, and $d_{i,t-1} = d_{i',t-1}$). However, instead of requiring the synthetic control market to have no branches, we require that it does not have a branch before year t but will have a branch opened in year $t+\tau$, where τ is equal to 10, 8, 6, 4, and 2 in columns (2), (3), (4), (5), and (6), respectively. In columns (7) to (11), we evaluate how the lagged effect of branch openings in control markets affect the DID estimates in the simulated data. Similar to what we do in column (1), each treatment market i in year t is matched with a synthetic control market i' of identical characteristics (i.e., $\Pi_{i,t-1} = \Pi_{i',t-1}$, $Z_i = Z_{i'}$, and $d_{i,t-1} = d_{i',t-1}$). However, instead of requiring the synthetic control market to have no branches, we require that it has a branch opened in year $t+\tau$, where τ is equal to -2, -4, -6, -8, and -10 in columns (7), (8), (9), (10), and (11), respectively.

Specifically, to obtain the estimates in column (1) of Table OA.12, for each year $t = 1987, ..., 1996$ in the simulated data, we first find all treatment markets (i.e., the markets in which a branch is opened in year t). Next, for each treatment market i in year t, we construct a synthetic control market i' with identical characteristics (i.e., $\Pi_{i,t-1} = \Pi_{i',t-1}$, $Z_i = Z_{i'}$, and $d_{i,t-1} = d_{i',t-1}$) in year $t-1$ prior to the treatment. We ensure that the synthetic control market does not have a bank branch opened before year t nor a branch opened after year t to ensure that our model-implied DID estimates are not confounded by the anticipation or lagged effects of branch openings in the synthetic control market. Because there are 431 treatment markets during the 1987-1996 period, in total, we create 431 synthetic control markets. Next, we simulate the outcome variables for each synthetic control market, given the equilibrium interest rates, wages, and branch locations in each year solved in our baseline model. Note that we solve and simulate the baseline model before we create the synthetic control markets. Thus, their creation does not affect the baseline model. The synthetic control markets are introduced solely for the purpose of conducting the DID estimation using the simulated data. Obviously, the synthetic control markets provide perfect controls for the treatment markets because they essentially capture what would happen if the treatment markets were not treated (i.e., if the branches were not opened in treatment markets in the simulated data). Based on the 431 matched pairs of treatment and synthetic control markets, we run the DID estimation in the simulated data according to specification (2.1) of the main text.

In columns (2) to (6) of Table OA.12, we evaluate how the anticipation effect of branch openings in control markets affects the DID estimates in the simulated data. Similar to our discussion above for column (1) of Table OA.12, each treatment market i in year t is matched with a synthetic control market i' with identical characteristics (i.e., $\Pi_{i,t-1} = \Pi_{i',t-1}$, $Z_i = Z_{i'}$, and $d_{i,t-1} = d_{i',t-1}$). However, instead of requiring the synthetic control market to have no branches, we require that it does not have a branch before year t but that it has a branch opened in year $t+\tau$, where τ is equal to 10, 8, 6, 4, and 2 in columns (2), (3), (4), (5), and (6), respectively. Comparing columns (1) and (6), we see that the estimated effects of branch openings are smaller if there are anticipated branch openings in the control markets in year $t+2$. However, the differences between columns (1) and (6) are very small, suggesting that the anticipation effect does not significantly affect the model-implied DID estimates in our simulated data. Moreover, as we increase τ from 2 to 10 (moving from column (6) to column (2)), the anticipation effect becomes even weaker and the estimates converge to the benchmark estimates in column (1). For example, the estimates in column (4), which correspond to the case with $\tau = 6$, are identical to those in column (1). Indeed, we show in Figure OA.52 that the anticipation of future branch openings will not have a substantial impact on the local market.

In columns (7) to (11) of Table OA.12, we evaluate how the lagged effect of branch openings in

control markets affects the DID estimates in the simulated data. Again, similar to the procedure described for column (1) of Table OA.12, each treatment market i in year t is matched with a synthetic control market i' with identical characteristics (i.e., $\Pi_{i,t-1} = \Pi_{i',t-1}$, $Z_i = Z_{i'}$, and $d_{i,t-1} = d_{i',t-1}$). However, instead of requiring the synthetic control market to have no branches, we require that it has a branch opened in year $t+\tau$, where τ is equal to -2, -4, -6, -8, and -10 in columns (7), (8), (9), (10), and (11), respectively. Comparing columns (1) and (7), we see that the estimated effects of branch openings are smaller if there are branch openings in the control markets in year $t-2$. The differences between columns (1) and (7) are very large, suggesting that the lagged effect significantly affects the model-implied DID estimates in our simulated data. Moreover, as we reduce τ from -2 to -10 (moving from column (7) to column (11)), the lagged effect becomes weaker and the estimates converge to the benchmark estimates in column (1). However, even if we focus on $\tau = -6$, meaning that branches in the matched control markets are opened 6 years before the treatment year for the treatment markets, the lagged effect still significantly affects the DID estimates. That is, the estimates in column (9) are different from those in column (1). Indeed, we show in Figure OA.52 that new branch openings have lasting effects on the local market because it takes time to converge to the new steady state due to the existence of financial constraints.

4 Supplemental Information for Quantitative Analyses

In this section of the appendix, we provide supplemental information for quantitative analyses. Subsection 4.1 details the calibration of migration costs and quantifies the role of migration in our model. Subsection 4.2 discusses the complementarity between the credit and deposit channels of bank expansion and conducts robustness checks for different values of parameter θ_ζ. Subsection 4.3 discusses the sensitivity of the main quantitative results to alternative model specifications and calibration. Subsection 4.4 presents the implications of an alternative model specification in which new branches opened between 1987 and 1996 are exogenously placed at the same locations as those in the data. Subsection 4.5 presents the implications of an alternative model specification in which new branch locations are randomly selected. Subsection 4.6 presents the model-implied equilibrium flow of funds across markets. Subsection 4.7 shows that our model-implied transitional dynamics for the 1986-1996 period are similar even if we allow new branches to open after 1996, because the anticipation effect of branch openings is small. In Subsection 4.8, we apply the model to analyze the post 1996 financial crisis.

4.1 Migration: Calibration and Quantitative Implications

Subsection 4.1.1 details the calibration of migration costs. Subsection 4.1.2 studies the quantitative implications when the model is calibrated to match a lower or a higher migration rate. Subsection 4.1.3 shows that the quantitative implications of migration are similar if we allow the fixed migration cost to increase with migration distance.

4.1.1 Calibration of Migration Costs

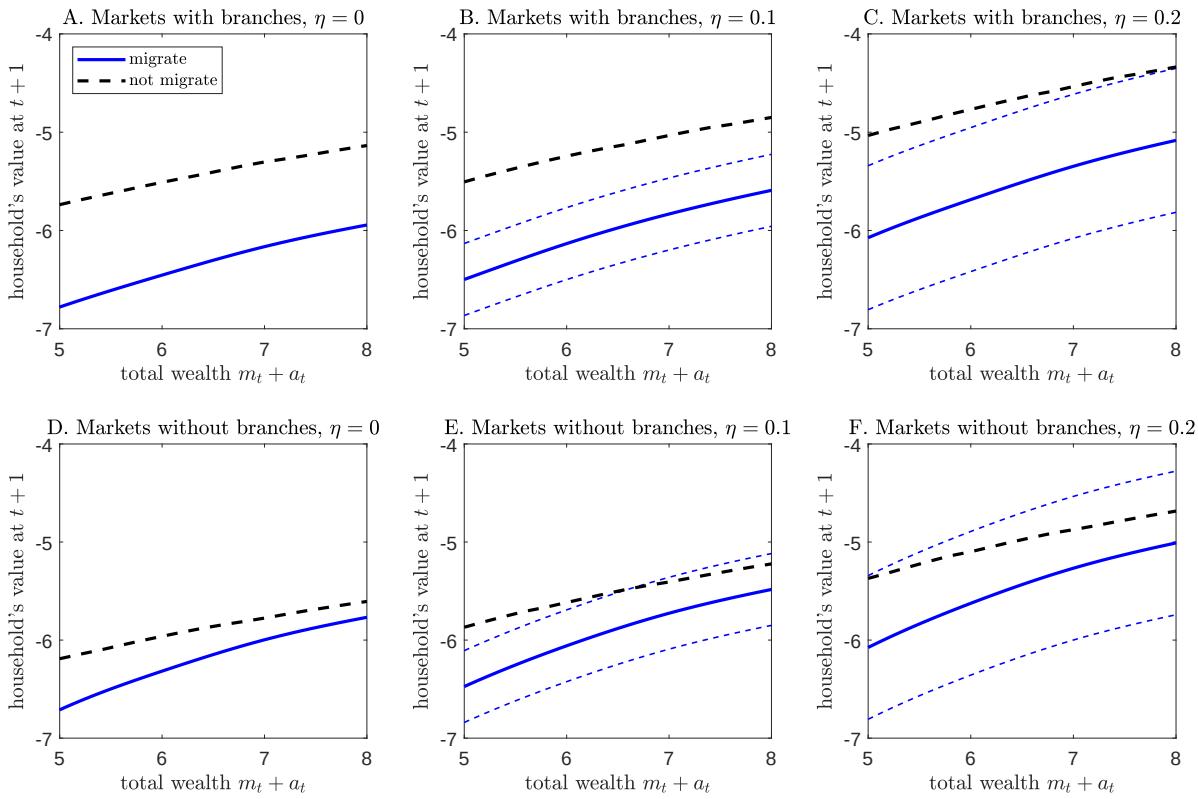
We discuss the identification of parameters κ and η. As we discuss in Section 4 of the main text, these two parameters are calibrated to match two moments, the overall migration rate from 1986 to 1996 and the ratio of the out-migrant share of markets with branches to that of markets without branches.

Role of η. Online Appendix 3.3.4 above shows that a lower κ will expand the space where households choose to migrate, increasing the equilibrium migration rate. Thus, it is clear that by varying the parameter κ alone, we can make the model-implied migration rate for 1986-1996 consistent with the data. Now, we illustrate the role played by the parameter η, which governs the importance of idiosyncratic taste shocks in driving households' migration decisions.

Consider a household currently living in a market with a bank branch at t (i.e., $d_{i,t} = 0$). We fix the household's talent level at $z_t = 1.75$ (the median value of entrepreneurial talents) and the deposit-wealth ratio at $a_t/(m_t + a_t) = 0.2$ (the median value of this ratio across markets). In panel A of Figure OA.22, we set $\eta = 0$. The black dashed line plots the household's value $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i)$ at $t+1$, if the household stays in the current market i, as a function of its current wealth $m_t + a_t$. The blue solid line plots the household's value $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j)$ at $t+1$, if it migrates to another market j, which is far away from bank branches ($d_{j,t} = 12.8$ minutes, corresponding to the median car travel time to the nearest bank branch among all markets without branches in 1986). The blue solid line is far below the black dashed line, suggesting that the household optimally chooses to stay in the current market.

In panel B, we set $\eta = 0.1$ to introduce idiosyncratic taste shocks.[21] In this case, to decide whether to migrate at $t + 1$, the household will compare $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i) + \epsilon_{i,t+1}$ and

[21] Our calibrated value of η is 0.008 (see Table 2 of the main text), much smaller than the values used in panels B and C of Figure OA.22, but it can still generate a significant out-migrant share for markets with bank branches. The reason is that in our model solution, migration decisions are made by comparing the continuation values across all 1,428 markets. Households may choose to migrate out of their current markets for idiosyncratic reasons, even with a small value of η, because idiosyncratic taste shocks are drawn independently across markets. However, in panels B and C of Figure OA.22, we only compare the continuation values of two markets, i and j. To illustrate the role of η more clearly, we need to set it at a much higher value than the value in our calibration.

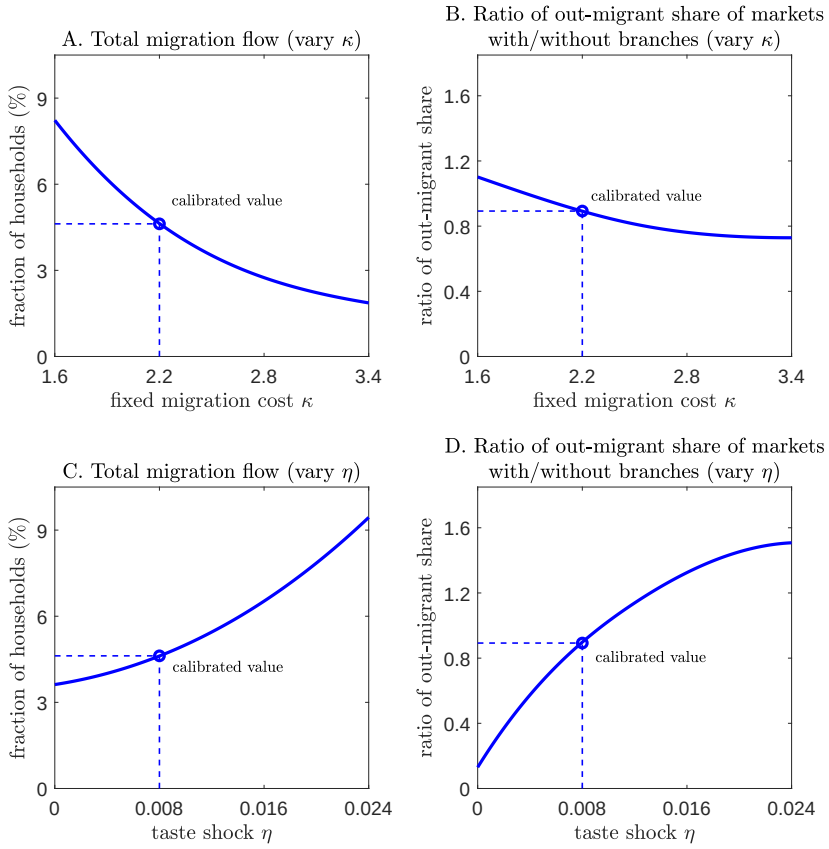


Note: Panels A to C illustrate the value function for a household living in a market with a bank branch ($d_{i,t} = 0$). Panels D to F illustrate the value function for a household living in a market far away from bank branches ($d_{i,t} = 12.8$ minutes, which corresponds to the median car travel time to the nearest bank branch among all markets without branches in 1986). In all panels, the x-axis is household total wealth $m_t + a_t$; and we fix household talent at $z_t = 1.75$, corresponding to the median value of entrepreneurs, and the deposit-wealth ratio at $a_t/(m_t + a_t) = 0.2$, corresponding to the median value across markets. In all panels, the black dashed lines represent the value of staying in the current market i, i.e. $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i)$. In panels A to C, the blue solid lines represent the value $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j)$ if the household migrates to market j that is far away from bank branches, with $d_{j,t} = 12.8$ minutes; and in panels D to F, the blue solid lines represent $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j)$ if the household migrates to market j that has a bank branch, i.e., $d_{j,t} = 0$. The blue dashed lines represent the 95% confidence interval of $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j) + \epsilon_{j,t+1} - \epsilon_{i,t+1}$. We set $\eta = 0$ in panels A and D; $\eta = 0.1$ in panels B and E; and $\eta = 0.2$ in panels C and F. The interest rate r_t and wages $\{w_{i,t}\}_{i=1}^{N}$ are set according to their equilibrium values in 1986. Other parameter values are set according to our calibration in Section 4 of the main text.

Figure OA.22: Value function of households in different markets.

$V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j) + \epsilon_{j,t+1}$.[22] To illustrate the effects of these idiosyncratic taste shocks, the blue dashed lines plot the 95% confidence interval of $\epsilon_{j,t+1} - \epsilon_{i,t+1}$ around the blue solid line. Clearly, the randomness generated by taste shocks makes the upper bound of the 95% confidence interval closer to the black dashed line. In panel C, we increase η to 0.2 and the 95% confidence interval expands further. Specifically, when total wealth $m_t + a_t = 7.9$, the upper bound of the

[22] A change in the value of η affects the values of $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, i)$ and $V_{t+1}(z_{t+1}, m_{t+1}, a_{t+1}, j)$. Thus, the blue solid and black dashed lines differ across panels A to C of Figure OA.22.



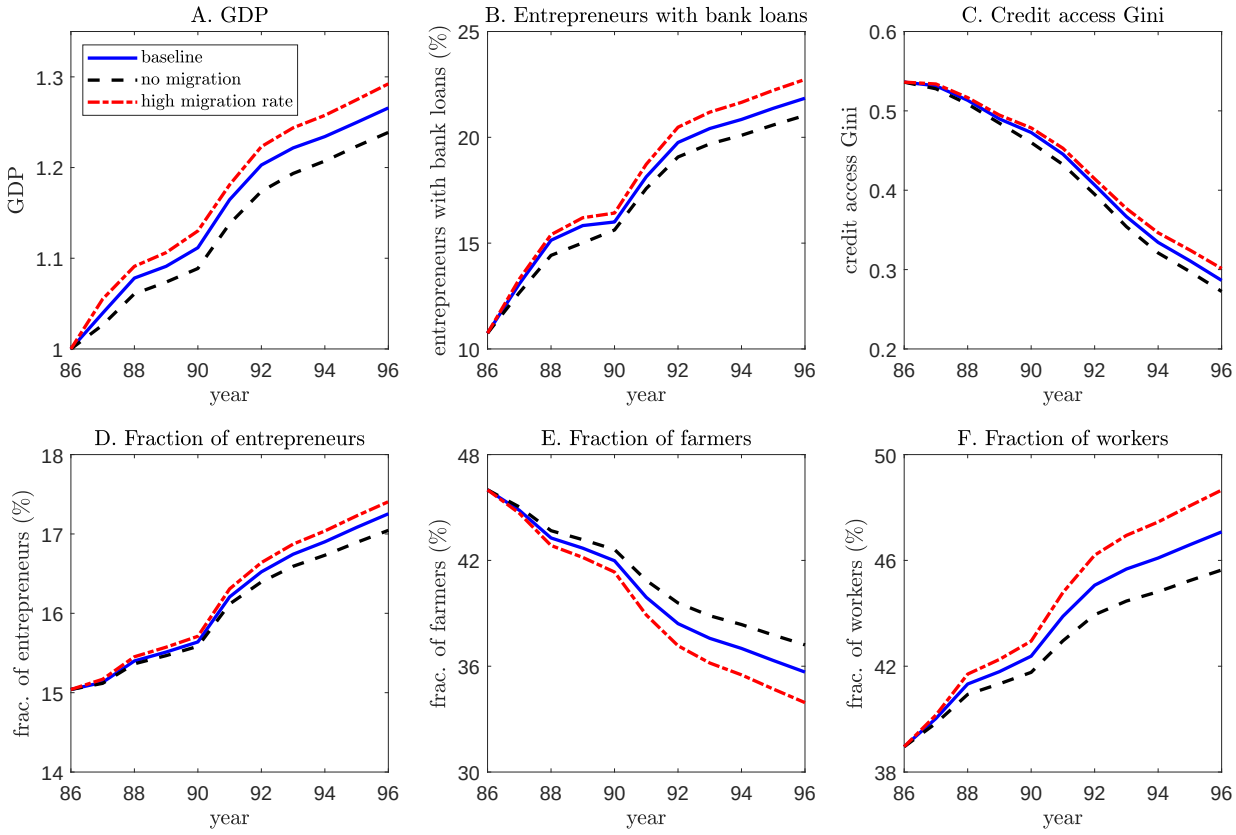
Note: This figure plots the model-implied migration rate from 1986 to 1996 and the ratio of the out-migrant share of markets with branches to that of markets without branches for different values of κ (panels A and B) and η (panels C and D), respectively. In panels A and B, we set $\eta = 0.008$ according to our calibration and vary κ along the x-axis. In panels C and D, we set $\kappa = 2.2$ according to our calibration and vary η along the x-axis. All other parameters are set according to our calibration in Table 2 of the main text.

Figure OA.23: The sensitivity of the migration rate and the ratio of the out-migrant share of markets with branches to that of markets without branches to parameters κ and η.

95% confidence interval crosses the black dashed line, suggesting that the household will choose to migrate to market j with a 2.5% probability.

Taken together, panels A to C of Figure OA.22 indicate that increasing η will lead to an increase in decisions to migrate for idiosyncratic reasons. This generates a larger percentage increase in the migration flow from markets with bank branches to markets without bank branches because such migration flows are rare, owing to the large gap in value attained by households between these two markets when $\eta = 0$.

Panels D to F of Figure OA.22 conduct the same experiments for a household of the same type, except that, the household currently lives in a market that is far from bank branches at t ($d_{i,t} = 12.8$ minutes). The black dashed lines plot the household's value if it stays in the current



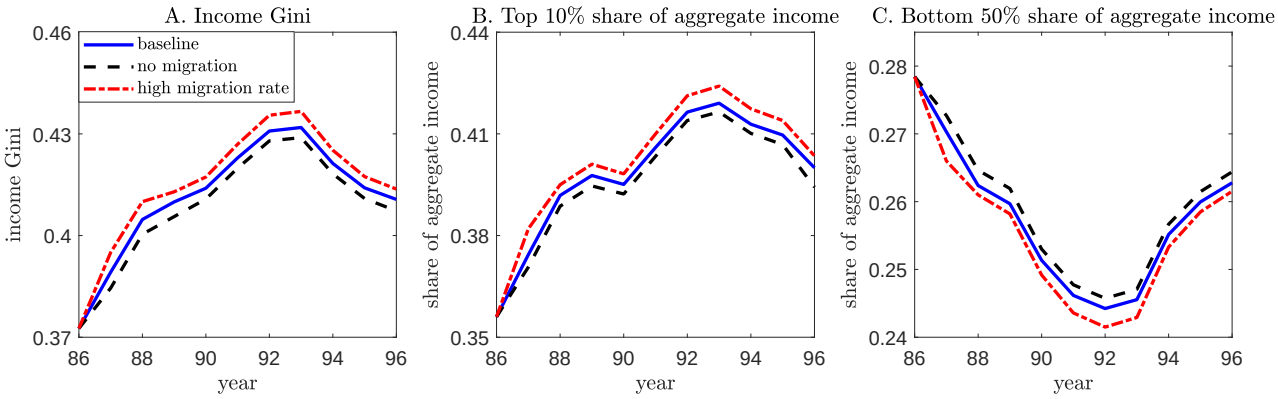
Note: The blue solid lines represent the aggregate dynamics in the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics under an alternative calibration with $\kappa = \infty$ and $\eta = 0$, which implies a zero migration rate. The red dash-dotted lines represent the aggregate dynamics under an alternative calibration with $\kappa = 1.2$ and $\eta = 0.005$, which implies a migration rate of 14.1%, twice higher than that in our baseline calibration, and a ratio of out-migrant shares of 0.89, similar to that in the baseline calibration. The other parameters are set at the same values as in our baseline calibration in Table 2 of the main text.

Figure OA.24: Different migration rates: Aggregate dynamics, 1986-1996.

market i, while the blue solid lines indicate its value if it migrates to another market j that has a local bank branch. Clearly, the gap between the two lines in panel D is much smaller than that in panel A, indicating that households living in markets without branches are more likely to migrate, as discussed in Online Appendix 3.3.4. When we increase η in panels E and F, household migration decisions are influenced more by idiosyncratic reasons.

Identification of Parameters κ and η. Figure OA.23 illustrates how the two parameters κ and η are identified by matching the two moments in the data on the 1986-1996 migration rate and the ratio of the out-migrant share of markets with branches to that of markets without branches.

In panels A and B of Figure OA.23, we set $\eta = 0.008$ according to our calibration in Table 2



Note: The blue solid lines represent the dynamics in the baseline model (which are the same as the blue solid lines in Figure 9 of the main text). The black dashed lines represent the dynamics under an alternative calibration with $\kappa = \infty$ and $\eta = 0$, which implies a zero migration rate. The red dash-dotted lines represent the dynamics under an alternative calibration with $\kappa = 1.2$ and $\eta = 0.005$, which implies a migration rate of 14.1%, twice higher than that in our baseline calibration, and a ratio of out-migrant shares of 0.89, similar to that in the baseline calibration. The other parameters are set at the same values as in our baseline calibration in Table 2 of the main text.

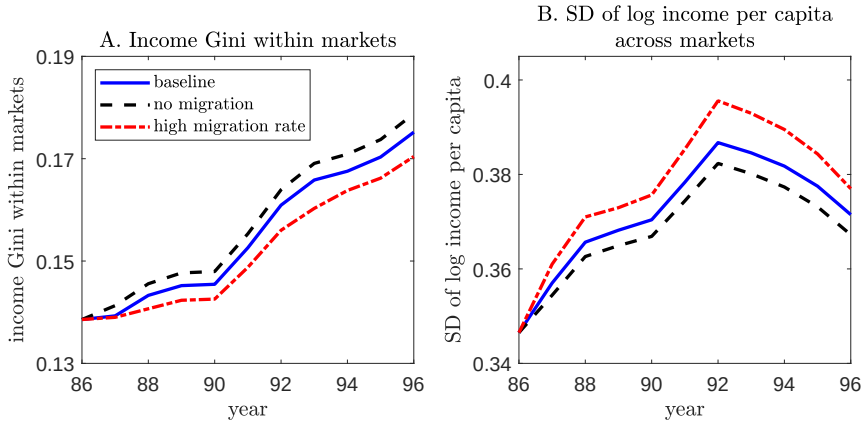
Figure OA.25: Different migration rates: Overall income inequality across households.

of the main text and vary κ along the x-axis. Panel A shows that the model-implied migration rate from 1986 to 1996 decreases monotonically as κ increases. Panel B shows that the ratio of the out-migrant share of markets with branches to that of markets without branches changes less significantly in comparison. Then, in panels C and D of Figure OA.23, we set $\kappa = 2.2$ according to our calibration in Table 2 of the main text and vary η along the x-axis. Both the model-implied migration rate and the ratio of the out-migrant share of markets with branches to that of markets without branches increase significantly as η increases. Therefore, Figure OA.23 clearly shows that the two moments have differential sensitivities to κ and η in our model, which provides the basis for the identification of these two parameters.

4.1.2 Quantitative Implications of Migration

To evaluate the role of migration, we solve a similarly calibrated model except for setting the fixed migration cost at $\kappa = \infty$ and the scale parameter of idiosyncratic taste shocks at $\eta = 0$. This alternative calibration essentially restricts households from migrating to other markets, resulting in a zero migration rate in equilibrium.

A comparison of our baseline model and the alternative model specification without migration flows indicates similar predicted distributions of bank branches in 1996. Of the 431 branches opened between 1987 and 1996, only two branch locations are different in the model without migration. Specifically, in the model without migration, there is one more branch opened in the northeast and south regions of Thailand, respectively, but the number of branches opened in the Bangkok metropolitan region is 28, compared with 30 in our baseline model, and 32 in the data.



Note: The blue solid lines represent the dynamics in the baseline model (which are the same as the blue solid lines in Figure 10 of the main text). The black dashed lines represent the dynamics under an alternative calibration with $\kappa = \infty$ and $\eta = 0$, which implies a zero migration rate. The red dash-dotted lines represent the dynamics under an alternative calibration with $\kappa = 1.2$ and $\eta = 0.005$, which implies a migration rate of 14.1%, twice higher than that in our baseline calibration, and a ratio of out-migrant shares of 0.89, similar to that in the baseline calibration. The other parameters are set at the same values as in our baseline calibration in Table 2 of the main text.

Figure OA.26: Different migration rates: Income inequality within and across markets.

Thus, by introducing migration, we marginally improve the predictions of bank branches around Bangkok.

The black dashed lines in Figure OA.24 plot the aggregate transitional dynamics in the alternative model specification without migration flows. Relative to our baseline model (blue solid lines), restricting households from migrating to other markets reduces GDP growth between 1986 and 1996 from 26.5% to 23.8%. This suggests that inter-market migration contributes 2.7% of GDP growth during this period in our baseline model. This additional GDP growth caused by migration is mainly due to talented but wealth constrained entrepreneurs migrating to markets with bank branches and producing more output as a result of obtaining loans, as indicated by Figure OA.8 in Online Appendix 3.3.4. Migration does not seem to play a quantitatively important role at the aggregate level because the model is calibrated to generate a low (4.7%) migration rate from 1986 to 1996, as in the data.

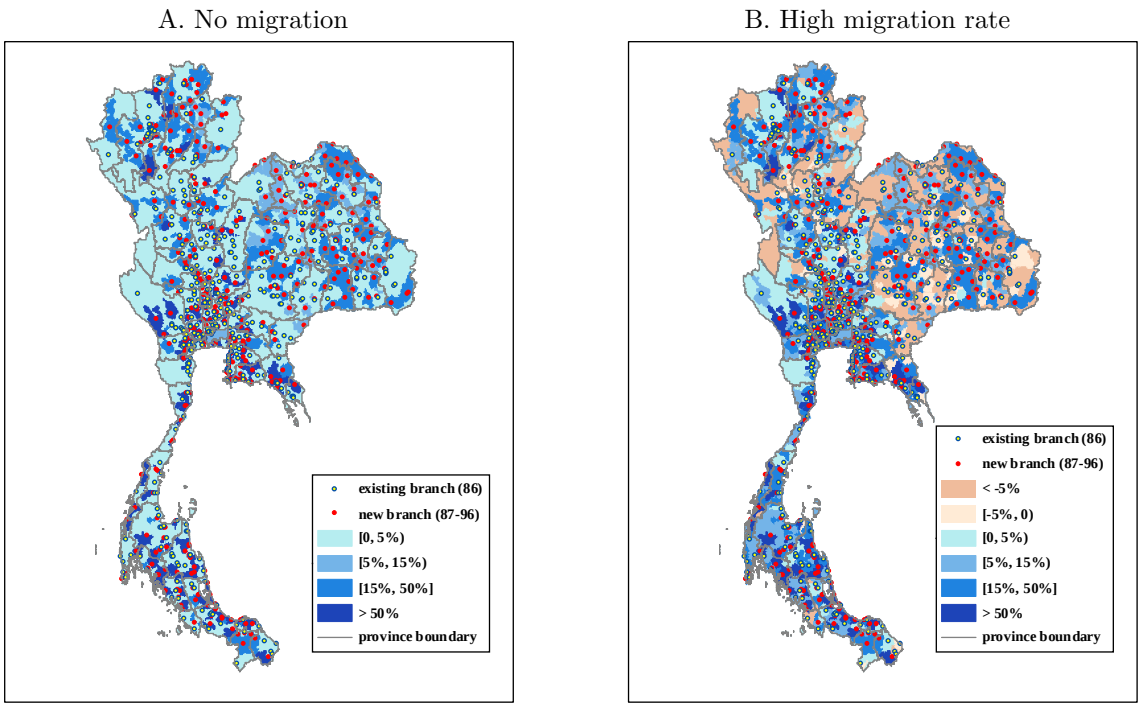
We perform a sensitivity analysis by calibrating the migration costs at $\kappa = 1.2$ and $\eta = 0.005$ to generate a migration rate of 14.1%, twice higher than that in our baseline calibration, and a ratio of out-migrant shares of 0.89, similar to that in the baseline calibration. The red dash-dotted lines in Figure OA.24 represent the model-implied dynamics under this calibration, which predicts GDP growth of 29.2% for the 1986-1996 period, indicating that migration contributes about 5.4% of GDP growth during this period, which is a much larger contribution than the one quantified under our baseline calibration. The labor transition from farming to wage labor also becomes more significant compared with our baseline calibration (panels E and F of Figure OA.24).

Figure OA.25 and panel B of Figure OA.26 show that our baseline model implies slightly higher overall income inequality and income differences across markets than the alternative model specification without migration flows. By contrast, panel A of Figure OA.26 shows that our baseline model implies slightly lower income inequality within markets than the alternative model specification without migration flows. As shown by the red dash-dotted lines in these figures, further increasing the migration rate to 14.1% would lead to a further increase in the level of overall and cross-market income inequality and a further decrease in the level of within-market income inequality. The main reason is that talented but wealth constrained entrepreneurs, who already have higher income than other households, migrate to markets with bank branches and, after obtaining access to bank loans, experience a further increase in their income. This will increase overall income inequality and income differences across markets. However, offsetting this, it leads to lower within-market income inequality because there is a greater concentration of talented people in the same market after migration, resulting in lower within-market talent and income differences.

Panel A of Figure OA.27 presents the welfare implications of the model in the absence of migration flows, which quantifies the welfare changes purely caused by bank expansion. The overall country-wide welfare increases by 19.0% from 1986 to 1996. In contrast to panel A of Figure 12 of the main text, all markets in panel A of Figure OA.27 have higher welfare in 1996 relative to their values in 1986 because there is no selection effect when households are not allowed to migrate across markets. The relative welfare gains across markets in panel A of Figure OA.27 are similar to those in panel A of Figure 12 of the main text. Markets initially far away from branches but receive new branch openings between 1987 and 1996 (red dots) experience large welfare gains. Markets that already have branches in 1986 (yellow dots) have small welfare gains due to the general equilibrium effect. Panel B of Figure OA.27 presents the welfare implications of the model when the calibrated migration rate is increased to 14.1%. The overall country-wide welfare increases by 21.0% from 1986 to 1996. The selection effect under this calibration is more significant than that in our baseline model due to the larger equilibrium migration flows. As a result, the market-level welfare changes are also more different across regions compared with panel A of Figure 12 of the main text.

4.1.3 Migration with Distance-Dependent Costs

In our baseline model, we assume that the fixed cost of migration κ is a constant. A more realistic specification is to allow κ to increase with migration distance. For example, Kennan and Walker (2011) allow the fixed cost κ to depend on migration distance, with the coefficient being identified by detailed location-to-location migration flow data. However, our CDD data do not provide information for location-to-location migration flows. As explained in Section 4 of the main text,



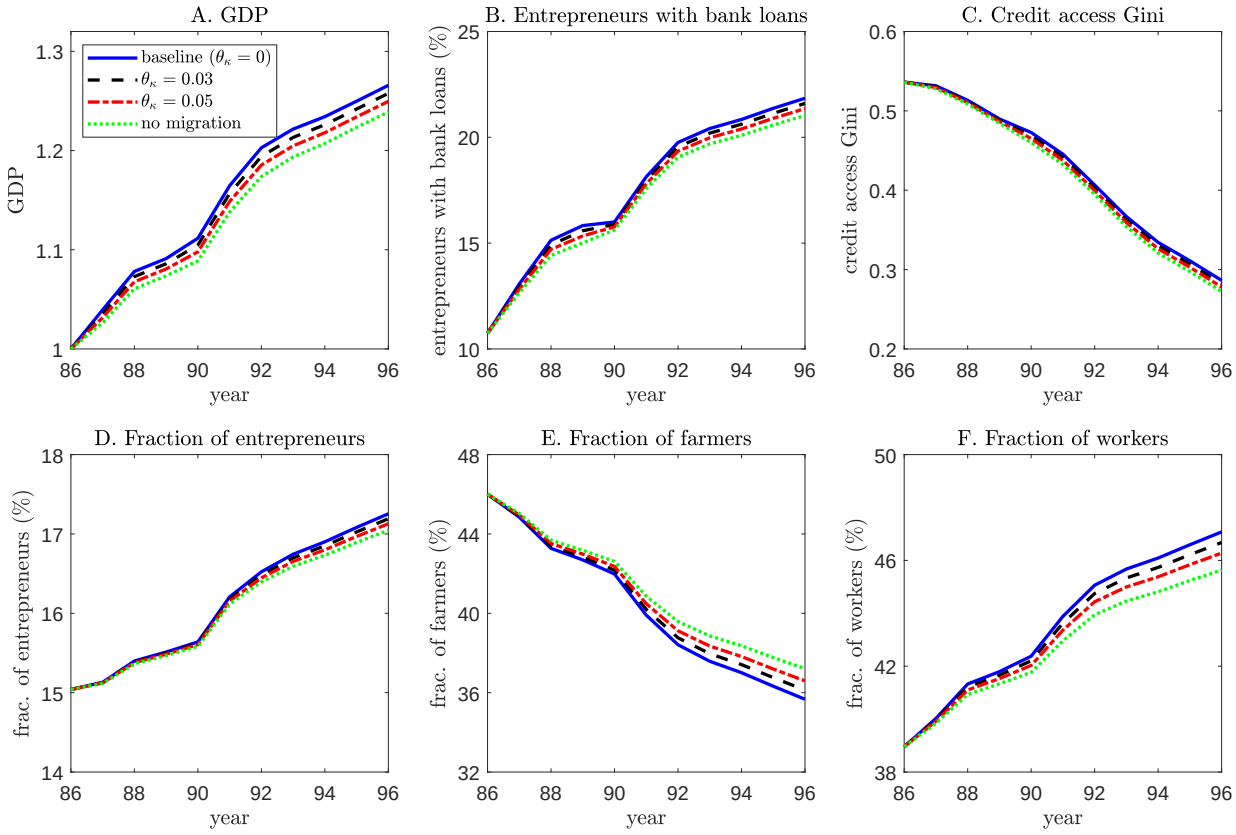
Note: This figure complements panel A of Figure 12 of the main text by presenting the cross-market welfare gains implied by two alternative calibrations for migration rates. Panel A plots the welfare gains under an alternative calibration with $\kappa = \infty$ and $\eta = 0$, which implies a zero migration rate. The overall country-wide welfare increases by 19.0% from 1986 to 1996. Panel B plots the welfare gains under an alternative calibration with $\kappa = 1.2$ and $\eta = 0.005$, which implies a migration rate of 14.1%, twice higher than that in our baseline calibration, and a ratio of out-migrant shares of 0.89, similar to that in the baseline calibration. The overall country-wide welfare increases by 21.0% from 1986 to 1996. The other parameters are set at the same values as in our baseline calibration in Table 2 of the main text.

Figure OA.27: Different migration rates: Market-level welfare gains, 1986-1996.

the CDD data only document the number of people working outside the township/tambon for each village in each survey year. Thus, we could not identify the sensitivity of migration flows to migration distance if we were to adopt the specification of Kennan and Walker (2011).

In this section of the appendix, we perform robustness checks to show that our model's quantitative implications are similar if we allow the fixed migration cost κ to depend on migration distance. The main reason is that the migration rate over the 1987-1996 period indicated by the data is not large. Once the model is calibrated to match the migration rate in the data, the contribution of migration flows to aggregate dynamics is quantified as small irrespective of how we specify the fixed migration cost.

Specifically, we follow Kennan and Walker (2011) and assume that the fixed migration cost linearly increases with migration distance, i.e., $\kappa_{ij} = \kappa + \theta_\kappa \tau_{ij}$ if households migrate from their current market i to market $j \neq i$. We consider two values for θ_κ by setting $\theta_\kappa = 0.03$ and



Note: The blue solid lines represent the aggregate dynamics in the baseline model (same as the blue solid lines in Figure 8 of the main text). The black dashed and red dash-dotted lines represent the aggregate dynamics in an alternative model specification in which the fixed migration cost linearly increases with distance, i.e., $\kappa_{ij} = \kappa + \theta_\kappa \tau_{ij}$ if households migrate from their current market i to market $j \neq i$. The black dashed and red dash-dotted lines correspond to $\theta_\kappa = 0.03$ and $\theta_\kappa = 0.05$, respectively. We calibrate κ and η to match the corresponding migration moments in panel B of Table 2 of the main text, as in the baseline model. The other parameters are set at the same values as in our baseline model. The green dotted lines represent the aggregate dynamics in an alternative model specification in which migration across markets is not allowed (achieved by setting $\kappa = \infty$ and $\eta = 0$), the same as the black dashed lines in Figure OA.24.

Figure OA.28: Migration with distance-dependent costs: Aggregate dynamics, 1986-1996.

$\theta_\kappa = 0.05$, respectively. The parameters κ and η are calibrated to match the corresponding migration moments in panel B of Table 2 of the main text, as in the baseline model. The other parameters are set at the same values as in our baseline model.

Figure OA.28 shows that compared with our baseline model (blue solid lines), allowing the fixed migration cost to depend on migration distance (black dashed and red dash-dotted lines) slightly reduces the predicted aggregate GDP growth. In addition, the trends in other aggregate variables change less dramatically over time compared with the baseline model, but the changes are more dramatic compared with the alternative specification in which migration across markets is not allowed (green dotted lines). Intuitively, making the migration cost increase with migration

distance implies that households will only focus on nearby markets when making migration decisions. This limits the set of markets to which households will migrate, which consequently reduces the output of those entrepreneurs who live far from branches but would like to migrate to markets with branches to borrow.

4.2 Credit Provision vs. Deposit Mobilization

In Subsection 4.2.1, we explain why there is a complementarity between the credit and deposit channels. In Subsection 4.2.2, we discuss the quantitative implications of our model for different values of parameter θ_ζ.

4.2.1 Complementarity Between the Credit and Deposit Channels

The complementarity between the deposit and credit channels (see panel A of Figure 8 of the main text) occurs for two reasons. First, the deposit channel allows households to access deposit accounts at lower costs $\zeta_{i,t}$, motivating households to hold a larger fraction of wealth in bank deposits. The interest paid by banks on deposits allows households to accumulate wealth faster than in the absence of such interest payments. Owing to collateral constraints, households will benefit from borrowing bank loans only when they have sufficient wealth to operate their businesses on a large scale. Thus, by facilitating wealth accumulation, deposit mobilization (i.e., a lower $\zeta_{i,t}$) allows talented but wealth constrained (potential) entrepreneurs to reap benefits from the lower credit entry costs $\psi_{i,t}$, amplifying the positive effect of the credit channel.

Second, deposit mobilization increases the supply of funds (i.e., loans to entrepreneurs), leading to a lower interest rate. In the absence of the credit channel, most markets face a high credit entry cost $\psi_{i,t}$; thus many talented entrepreneurs do not have access to credit. As a result, most of the additional funds provided through deposit mobilization are absorbed by less talented entrepreneurs located in markets that already had branches in 1986, and the funds exert a limited impact on aggregate output. By contrast, with the credit channel, most markets face a low $\psi_{i,t}$ after bank expansion, which allows talented entrepreneurs to borrow more and expand their businesses in response to a lower interest rate, improving the allocation of the additional funds. This intuition can be seen clearly in Figure OA.29. In panel A, the black dashed line represents the no-bank-expansion counterfactual (i.e., no new branches are opened between 1987 and 1996), where the upward trend reflects the effect of migration as talented entrepreneurs migrate to markets with branches and produce more output. Panel A shows that relative to the counterfactual of no bank expansion, the average talent (weighted by output) of entrepreneurs decreases under the deposit-channel counterfactual (i.e., only $\zeta_{i,t}$ decreases with new branch openings, blue solid line), indicating that additional funds are borrowed by entrepreneurs with lower talent levels. By contrast, panel B shows that relative to the credit-channel counterfactual

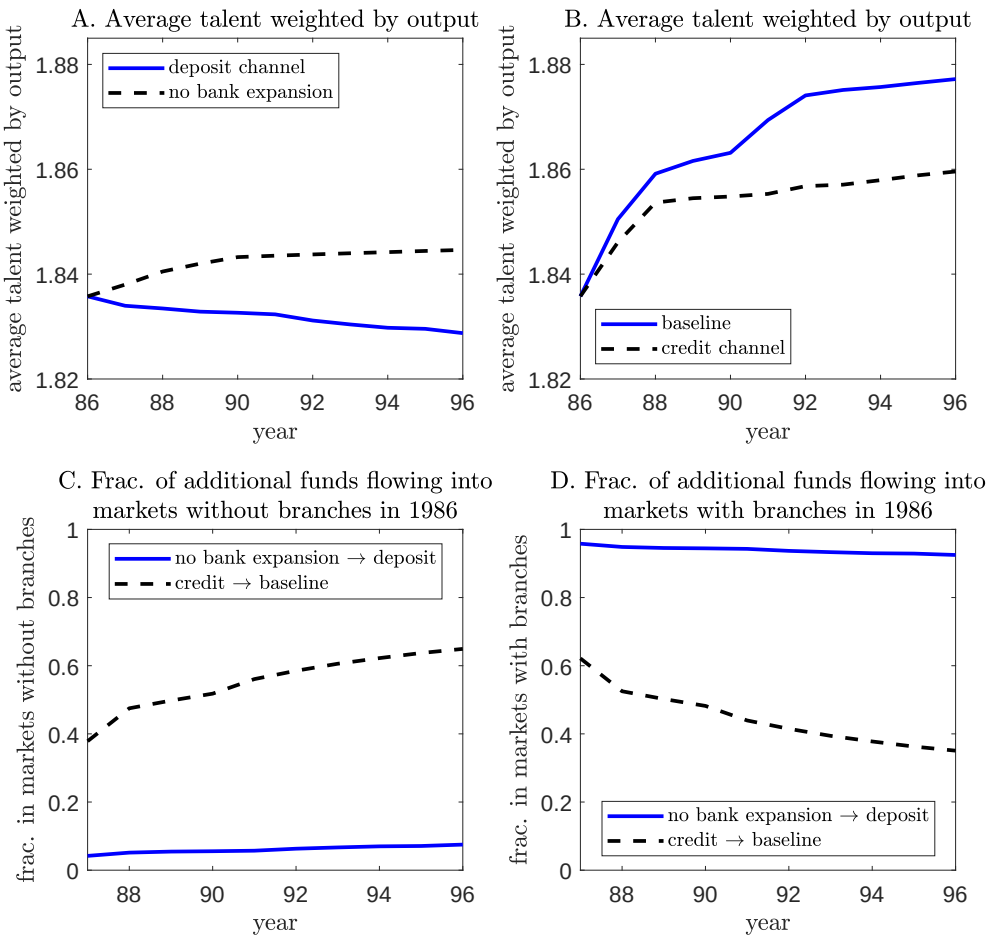


Figure OA.29: Complementarity between credit provision and deposit mobilization.

(i.e., only $\psi_{i,t}$ decreases with new branch openings, black dashed line), the average talent of entrepreneurs increases more in the baseline model (i.e., both $\zeta_{i,t}$ and $\psi_{i,t}$ decrease with new branch openings, blue solid line), indicating that additional funds are borrowed by entrepreneurs with higher talents.

We further show that the additional supply of funds generated by the deposit channel flows to households in different regions. Panels C and D of Figure OA.29 plot the fraction of additional funds flowing into markets without branches and markets with branches in 1986, respectively. Here, the additional funds associated with the blue solid lines refer to the increase in the total amount of household deposits when the economy moves from the no-bank-expansion to the deposit-channel counterfactual. It captures the increase in the total amount of household deposits due to the deposit channel in the absence of the credit channel (i.e., we assume that the credit entry cost $\psi_{i,t}$ remains unchanged with new branch openings). The blue solid lines show that

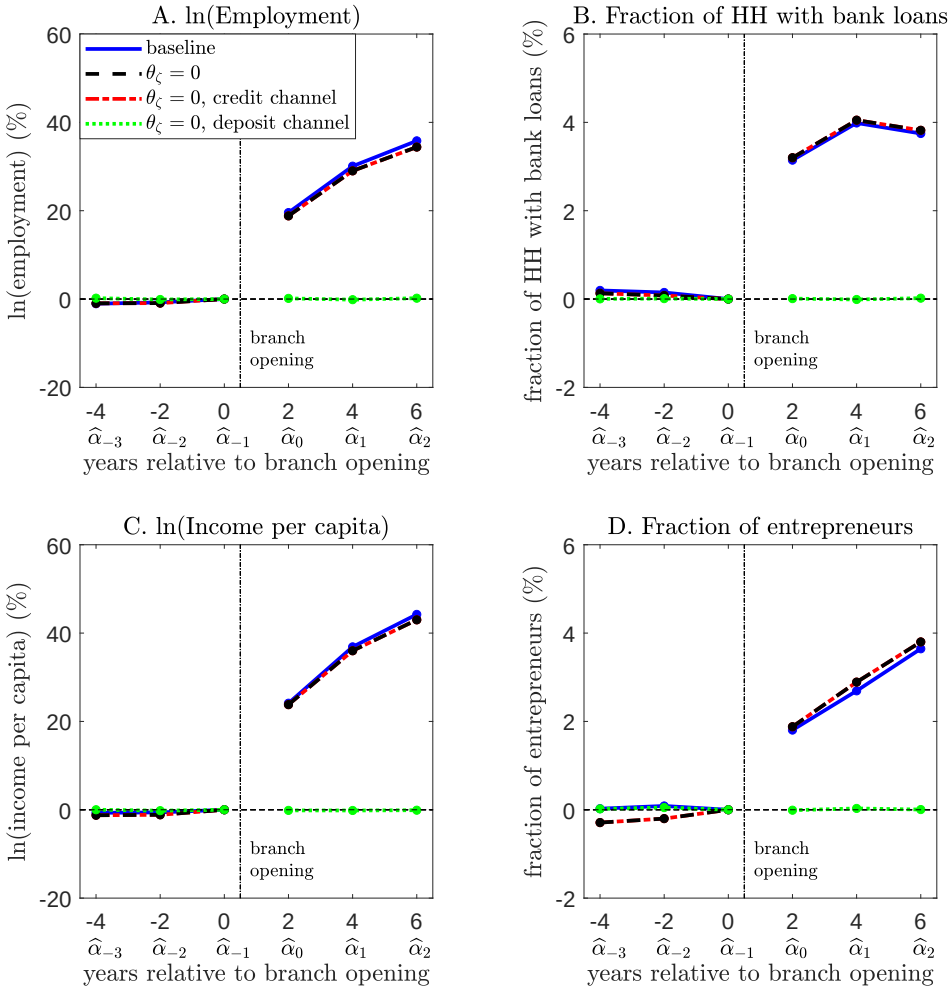
only a small fraction of additional funds flows into markets without branches in 1986 (panel C), whereas a large fraction flows into markets with branches (panel D). Intuitively, despite bank expansion, the credit entry cost $\psi_{i,t}$ remains high in markets without branches in 1986, because new branch openings only reduce the portfolio adjustment cost $\zeta_{i,t}$ in these markets. Thus, facing a high credit entry cost, entrepreneurs in these markets do not have the capacity to borrow additional funds, which are consequently absorbed by less talented entrepreneurs in markets with branches in 1986.

By contrast, the additional funds associated with the black dashed lines refer to the increase in total household deposits when the economy moves from the credit-channel counterfactual to the baseline model. It captures the increase in the total amount of household deposits due to the deposit channel (i.e., the portfolio adjustment cost $\zeta_{i,t}$ decreases with new branch openings) in the presence of the credit channel (i.e., the credit entry cost $\psi_{i,t}$ decreases with new branch openings). Compared with the blue solid lines, the black dashed lines show that a much larger fraction of additional funds flows into markets without branches in 1986 (panel C), whereas a much smaller fraction flows into markets with branches in 1986 (panel D). Intuitively, the credit entry cost $\psi_{i,t}$ decreases in markets without branches in 1986 after bank expansion because new branch openings reduce both the portfolio adjustment cost $\zeta_{i,t}$ and the credit entry cost $\psi_{i,t}$ in these markets. The lower credit entry cost in these markets allows talented local entrepreneurs to borrow additional funds.

4.2.2 Sensitivity Analyses for Parameter θ_ζ

In this section of the appendix, we present the implications of bank expansion for different values of parameter θ_ζ, which governs the sensitivity of portfolio adjustment costs $\zeta_{i,t}$ to the distance from the nearest bank branch $d_{i,t}$. We have two main findings. First, calibrating a lower (higher) sensitivity θ_ζ will decrease (increase) the quantitative importance of the deposit channel relative to the credit channel of bank expansion. Second, the calibrated value of θ_ζ has little effect on the aggregate implications of bank expansion (incorporating both channels). The main reason is that the aggregate implications of bank expansion are largely determined by the market-level effects of branch openings, which are disciplined to match the empirical DID estimates under our calibration. That is, for some arbitrary choice for the value of θ_ζ, we can always calibrate the two key parameters, θ_ψ and ξ, along with other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, and γ) to match the same moments presented in panel B of Table 2 in the main text. Because the DID estimates for the market-level effects of branch openings are matched by recalibrating θ_ψ and ξ, the model will yield similar aggregate implications of bank expansion.

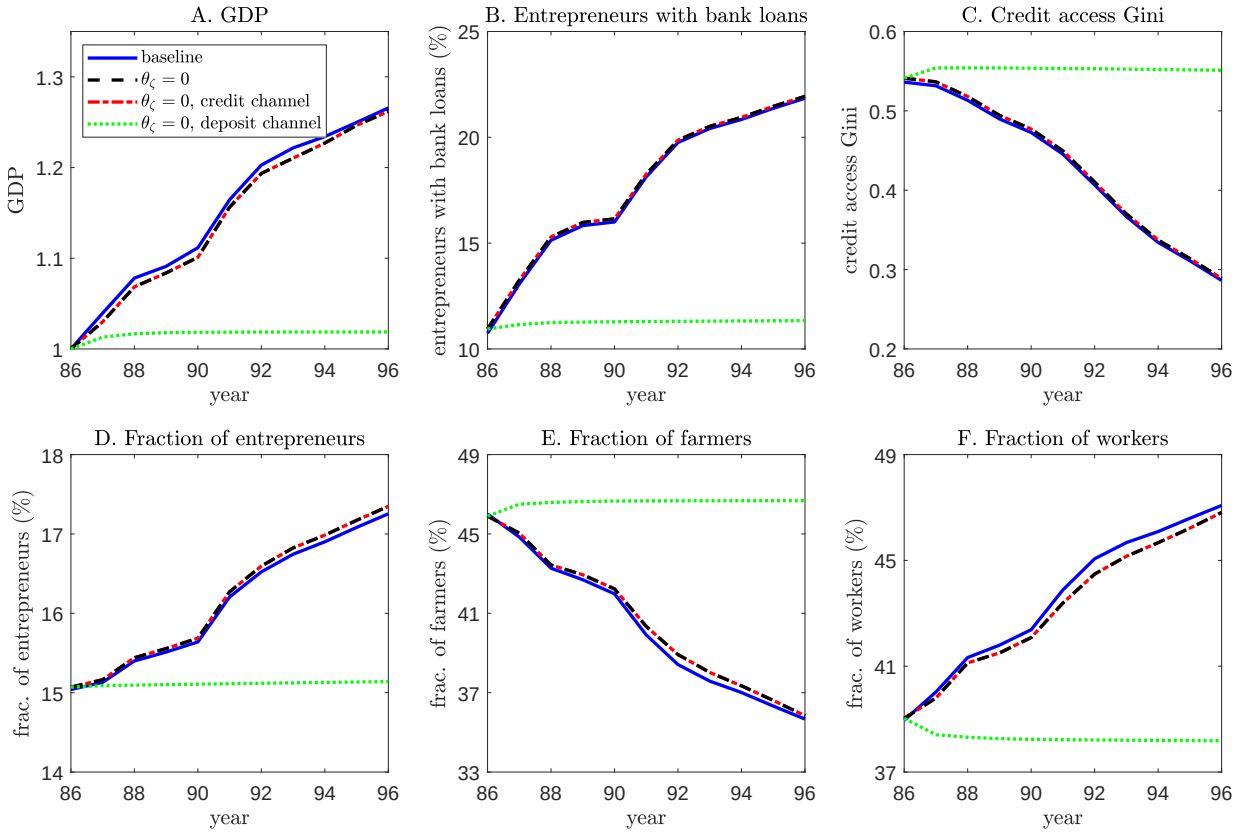
First, we consider an alternative calibration with $\theta_\zeta = 0$, implying that the distance from the nearest bank branch $d_{i,t}$ has no effect on the portfolio adjustment costs $\zeta_{i,t}$. We recalibrate other



Note: The blue solid lines represent the DID estimates in the baseline model (which are the same as the blue solid lines in Figure 5 of the main text). The black dashed lines represent the DID estimates under an alternative calibration with $\theta_\zeta = 0$. Other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, θ_ψ, and ξ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model. Specifically, the two key parameters are calibrated at $\theta_\psi = 0.25$ and $\xi = 0.59$. The red dash-dotted and green dotted lines represent the impacts through the credit and deposit channels, respectively, under this alternative calibration.

Figure OA.30: $\theta_\zeta = 0$: Market-level effects of branch openings.

parameters so that the moments reported in Table 2 in the main text are matched, except for the sensitivity of the deposit-cash ratio to $d_{i,t}$, which is matched by calibrating θ_ζ under our baseline calibration. The black dashed lines in Figure OA.30 present the model-implied market-level impacts of branch openings under this alternative calibration with $\theta_\zeta = 0$, which are very similar to the impacts implied by our baseline model (blue solid lines) because both calibrations target the same moments. Under this alternative calibration with $\theta_\zeta = 0$, $\zeta_{i,t}$ no longer depends on $d_{i,t}$. Thus, by definition, the deposit channel has no market-level effect, as shown by the flat green

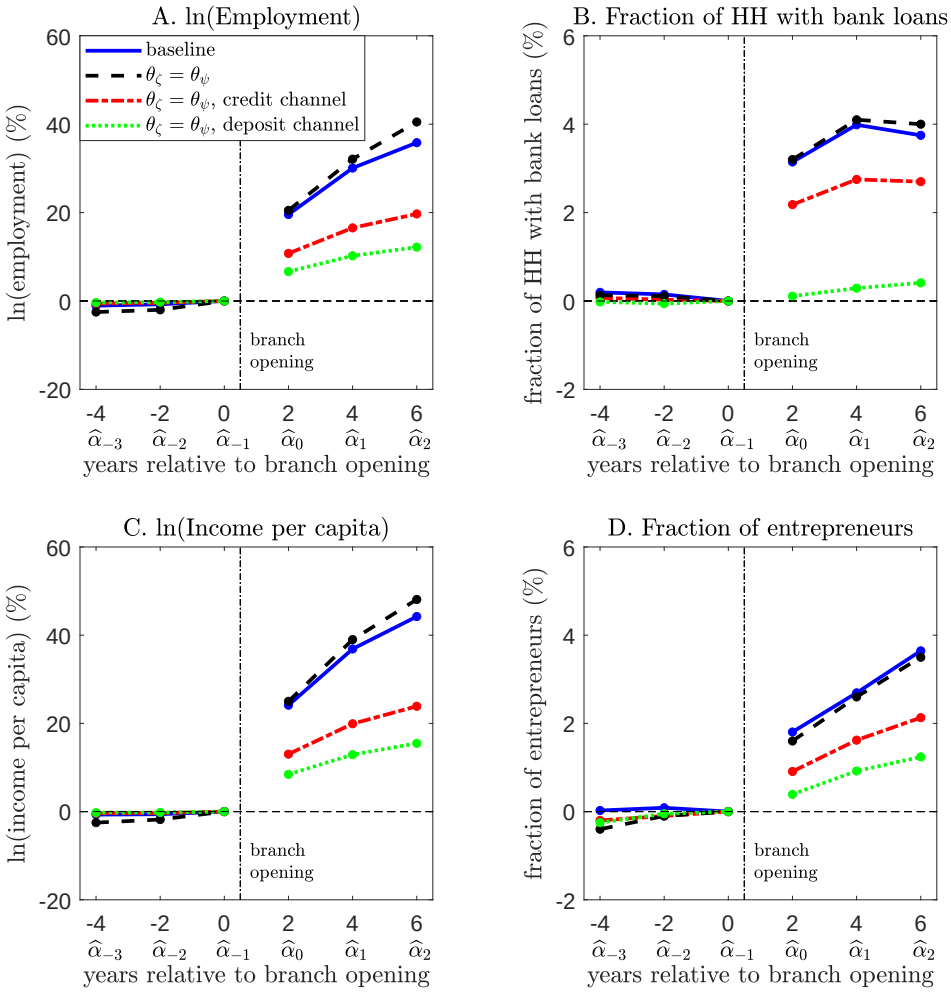


Note: The blue solid lines represent the aggregate dynamics in the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics under an alternative calibration with $\theta_\zeta = 0$. Other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, θ_ψ, and ξ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model. Specifically, the two key parameters are calibrated at $\theta_\psi = 0.25$ and $\xi = 0.59$. The red dash-dotted and green dotted lines represent the impacts through the credit and deposit channels, respectively, under this alternative calibration.

Figure OA.31: $\theta_\zeta = 0$: Aggregate dynamics, 1986-1996.

dotted lines in Figure OA.30. Moreover, the red dash-dotted lines overlap with the black dashed lines in Figure OA.30, indicating that the market-level impact of branch openings is caused purely by the credit channel.

The black dashed lines in Figure OA.31 present the model-implied aggregate dynamics under this alternative calibration with $\theta_\zeta = 0$. Not surprisingly, the aggregate dynamics are very similar to those implied by our baseline model (blue solid lines), given the similar DID estimates for the market-level impacts of branch openings shown in Figure OA.30. Specifically, GDP growth between 1986 and 1996 is 26.1% (the black dashed line in panel A) under this alternative calibration with $\theta_\zeta = 0$, which is very similar to the 26.5% implied by our baseline model (the blue solid line in panel A). The credit channel (red dash-dotted lines) implies identical aggregate dynamics to those of the black dashed lines, for reasons discussed above. However, in the deposit-

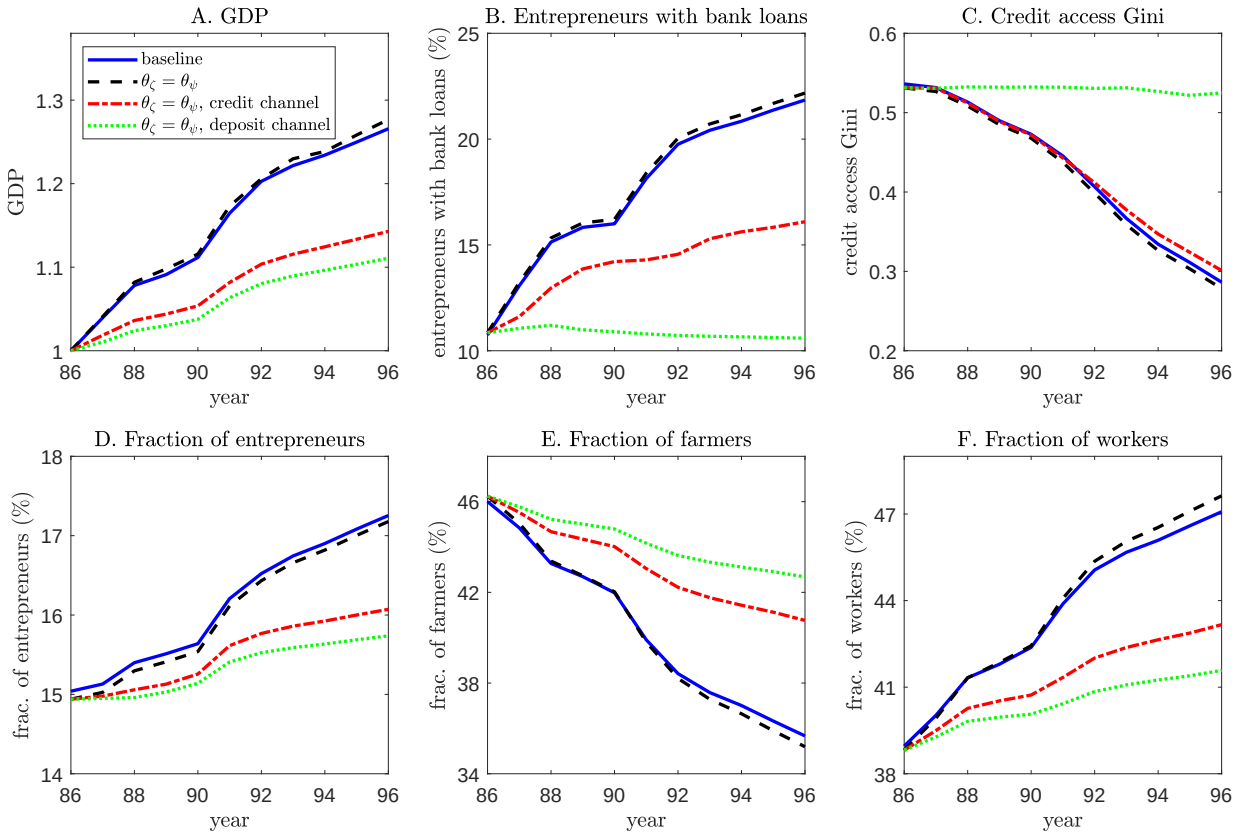


Note: The blue solid lines represent the DID estimates in the baseline model (which are the same as the blue solid lines in Figure 5 of the main text). The black dashed lines represent the DID estimates under an alternative calibration with $\theta_\zeta = \theta_\psi$. Other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, θ_ψ, and ξ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model. Specifically, the two key parameters are calibrated at $\theta_\psi = 0.17$ and $\xi = 0.59$. The red dash-dotted and green dotted lines represent the impacts through the credit and deposit channels, respectively, under this alternative calibration.

Figure OA.32: $\theta_\zeta = \theta_\psi$: Market-level effects of branch openings.

channel counterfactual (green dotted lines), there is a slight GDP growth of 1.9% between 1986 and 1996 due to inter-market migration, which has nothing to do with the deposit channel of bank expansion itself (see Figure OA.31, which indicates that the deposit channel has no market-level effects).

Second, we consider an alternative calibration by assuming that $\theta_\zeta = \theta_\psi$, which implies that the distance $d_{i,t}$ has similar effects on the credit entry cost $\psi_{i,t}$ and the portfolio adjustment cost $\zeta_{i,t}$. As we explain in Section 3.2 of the main text, $\psi_{i,t}$ reflects both transportation costs and

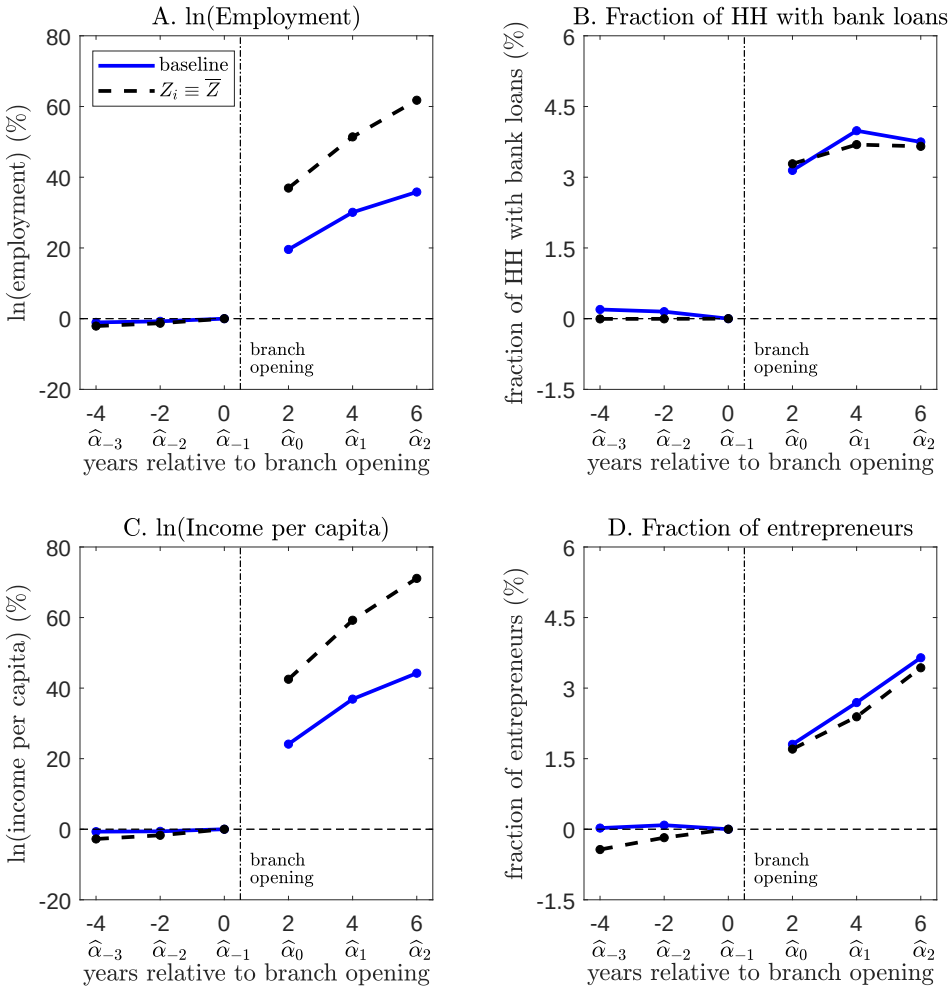


Note: The blue solid lines represent the aggregate dynamics in the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics under an alternative calibration with $\theta_\zeta = \theta_\psi$. Other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, θ_ψ, and ξ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model. Specifically, the two key parameters are calibrated at $\theta_\psi = 0.17$ and $\xi = 0.59$. The red dash-dotted and green dotted lines represent the impacts through the credit and deposit channels, respectively, under this alternative calibration.

Figure OA.33: $\theta_\zeta = \theta_\psi$: Aggregate dynamics, 1986-1996.

information costs, whereas $\zeta_{i,t}$ may only reflect transportation costs, implying that $\psi_{i,t}$ is more sensitive than $\zeta_{i,t}$ is to $d_{i,t}$. Thus, setting $\theta_\zeta = \theta_\psi$ can arguably be considered an upper-bound calibration for θ_ζ. We recalibrate other parameters so that the moments reported in Table 2 of the main text are matched, except for the sensitivity of the deposit-cash ratio to $d_{i,t}$ which is matched by calibrating θ_ζ under our baseline calibration. Specifically, θ_ψ $(= \theta_\zeta)$ is calibrated at 0.17 to match the DID estimate (α_0) for credit access in Table 2 in the main text.

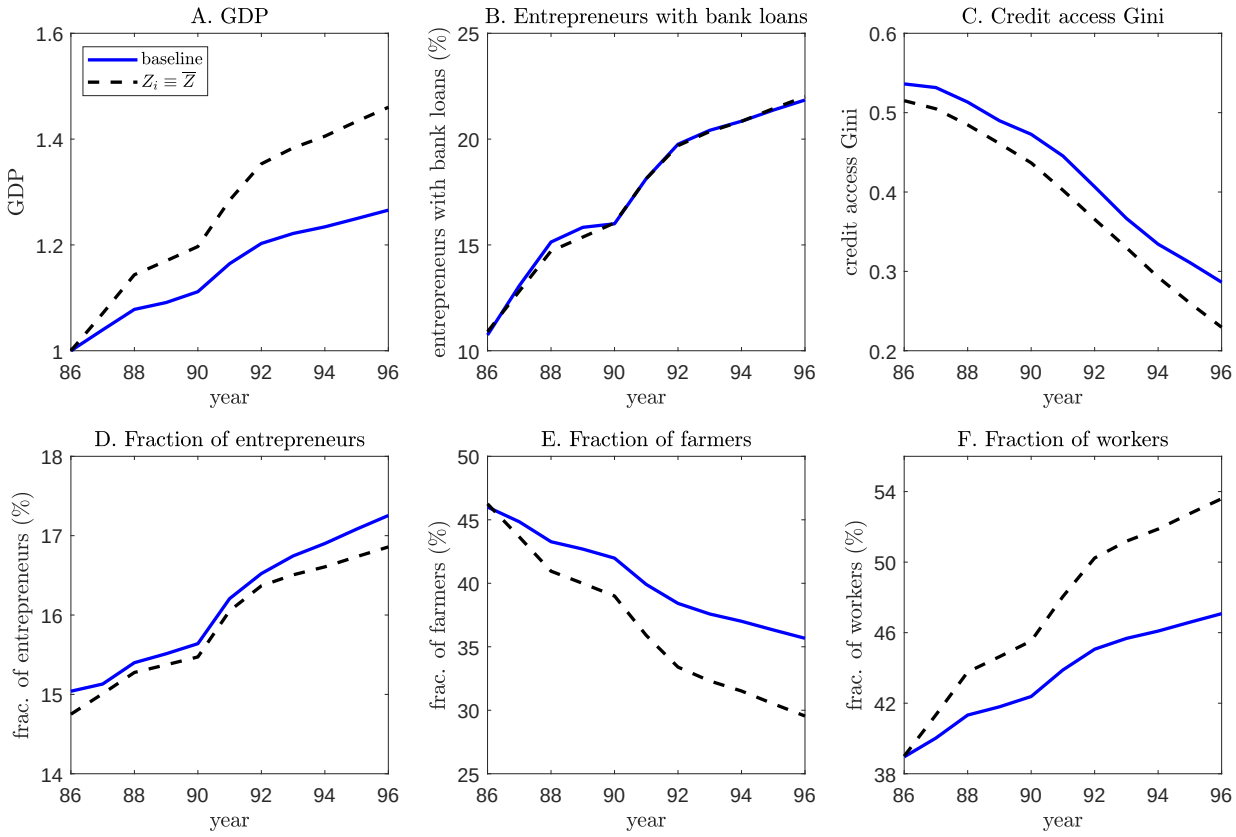
Figures OA.32 and OA.33 present the market-level impacts of branch openings and the aggregate implications of bank expansion, respectively, implied by our model under this alternative specification of $\theta_\zeta = \theta_\psi$. Panel A of Figure OA.33 shows that GDP growth for the 1986-1996 period is 27.6% (the black dashed line in panel A) under this alternative calibration with $\theta_\zeta = \theta_\psi$, which is very similar to the 26.5% implied by our baseline model (the blue solid line in panel A).



Note: The blue solid lines represent the DID estimates in the baseline model (which are the same as the blue solid lines in Figure 5 of the main text). The black dashed lines represent the DID estimates in an alternative model specification in which all markets have the same productivity (achieved by setting $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$). In this alternative model specification, instead of recalibrating the two parameters θ_ψ and ξ to match the DID estimates for credit access and employment, we set $\theta_\psi = 0.215$ and $\xi = 0.59$ at the same values as in our baseline calibration for comparison purposes. The other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model.

Figure OA.34: Productivity heterogeneity across markets: Market-level effects of branch openings.

Of the 27.6% GDP growth, the credit channel (red dash-dotted line) contributes 14.2% and the deposit channel (green dotted line) contributes 10.7%. By contrast, under our baseline calibration (panel F of Figure 8 in the main text), these two channels generate GDP growth of 17.8% and 6.5%, respectively. Thus, increasing θ_ζ to the same value as θ_ψ does not change the overall effect of bank expansion on GDP growth greatly, but it increases the relative importance of the deposit channel, though the deposit channel remains less important than the credit channel.

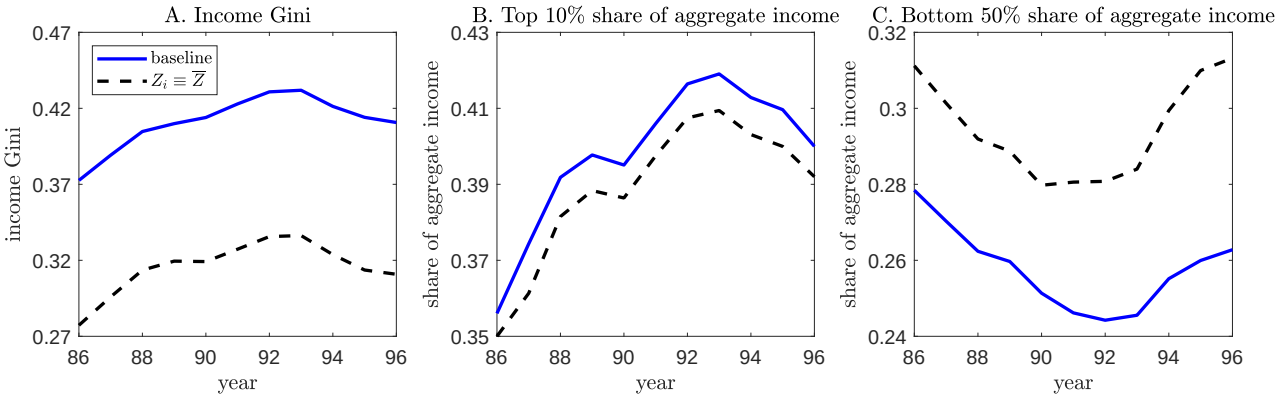


Note: The blue solid lines represent the aggregate dynamics in the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics in an alternative model specification in which all markets have the same productivity (achieved by setting $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$). This alternative model specification is calibrated according to the note for Figure OA.34.

Figure OA.35: Productivity heterogeneity across markets: Aggregate dynamics, 1986-1996.

4.3 Alternative Model Specifications and Calibration

In this appendix section, we discuss the sensitivity of the main quantitative results to alternative model specifications and calibration. Subsection 4.3.1 studies the model's quantitative implications when markets have the same level of productivity. Subsection 4.3.2 studies the model's quantitative implications when financial frictions also depend on the area of the market. Subsection 4.3.3 performs a sensitivity analysis for different values of ξ and θ_ψ, which govern the tightness of borrowing constraints and the dependence of financial frictions on the distance from the nearest bank branch, respectively. Subsection 4.3.4 performs a sensitivity analysis for different values of ρ and γ, which govern the stochastic process of individual talent. Finally, subsection 4.3.5 extends the baseline model with international capital flows.



Note: The blue solid lines represent the dynamics in the baseline model (which are the same as the blue solid lines in Figure 9 of the main text). The black dashed lines represent the dynamics in an alternative model specification in which all markets have the same productivity (achieved by setting $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$). This alternative model specification is calibrated according to the note for Figure OA.34.

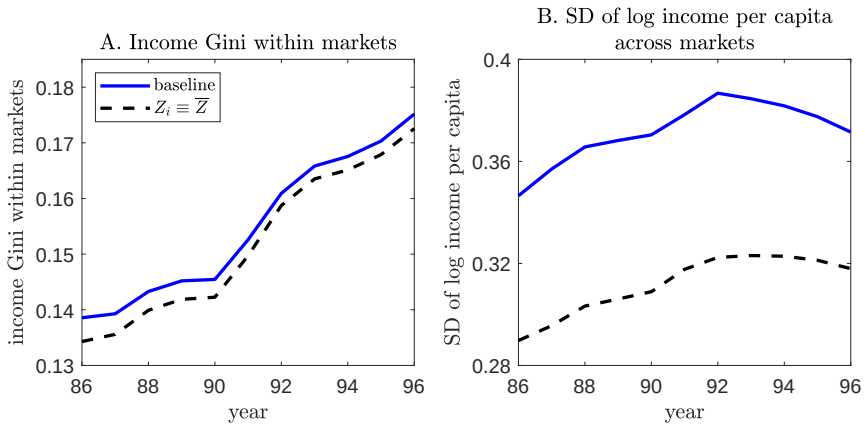
Figure OA.36: Productivity heterogeneity across markets: Overall income inequality across households.

4.3.1 Productivity Heterogeneity Across Markets

Section 5.2 of the main text shows that the productivity difference across markets plays an important role in predicting the branch locations in the data, especially for the Bangkok metropolitan and northeast regions. In this section of the appendix, we evaluate the role of cross-market productivity heterogeneity in determining the impacts of branch expansion.

Specifically, we solve and calibrate an alternative model specification in which all markets have the same productivity, equal to the national mean, i.e., $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$. Because the two parameters θ_ψ and ξ crucially determine the quantitative impacts of bank expansion, for comparison with our baseline model (with productivity heterogeneity), we set $\theta_\psi = 0.215$ and $\xi = 0.59$ at the same values as in our baseline calibration. The other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model.

The black dashed lines in Figure OA.34 plot the DID estimates from the simulated data based on this alternative model specification without productivity heterogeneity. Compared with the baseline model (blue solid lines), eliminating productivity heterogeneity leads to significantly larger impacts of branch openings on local employment (panel A) and income per capita (panel C) in our simulated data, though the impacts on the fraction of households with bank loans (panel B) and the fraction of entrepreneurs (panel D) are similar. Consistent with the magnitudes of these market-level estimates, in Figure OA.35, the black dashed line in panel A shows that GDP growth between 1986 and 1996 is 46.1% in the absence of productivity heterogeneity across markets; similarly, the black dashed line in panel F shows that the fraction of workers would increase to



Note: The blue solid lines represent the dynamics in the baseline model (which are the same as the blue solid lines in Figure 10 of the main text). The black dashed lines represent the dynamics in an alternative model specification in which all markets have the same productivity (achieved by setting $Z_i \equiv \overline{Z} = 1$ for all $i \in \mathcal{C}$). This alternative model specification is calibrated according to the note for Figure OA.34.

Figure OA.37: Productivity heterogeneity across markets: Income inequality within and across markets.

54.0% in 1996 in the absence of productivity heterogeneity across markets. Both estimates are much larger than those implied by our baseline model with productivity heterogeneity.

There are two reasons why introducing market-level productivity heterogeneity leads to smaller impacts of branch openings on employment and income per capita. First, note that our estimates indicate that markets with bank branches in 1986 tend to be more productive than other markets (see Figure OA.4). For example, the average productivity is 1.06 among the 406 markets with a branch in 1986, but it is 1.01 among the remaining 1,022 markets. Thus, the economy's GDP in 1986 is higher when markets exhibit productivity heterogeneity than when they do not. In part, this is simply because distributing the initial branches in 1986 to more productive markets leads to a higher aggregate GDP (see the red dash-dotted line in panel E of Figure OA.12). However, it is also due to calibrating a lower value of subsistence income f,[23] which increases the level of output in every market with farmers in 1986. Thus, in the baseline model with productivity heterogeneity, the higher base GDP in 1986 implies that the new branches opened between 1987 and 1996 will lead to a smaller percentage increase in GDP.

Second, in our baseline model with productivity heterogeneity, the central authority chooses to prioritize opening branches in more productive markets during the 1987-1996 period (see our discussions in Online Appendix 3.7.2). Owing to their higher productivity and the lower calibrated

[23]Holding parameter values unchanged, introducing productivity heterogeneity increases the fraction of farmers within the economy in 1986. The main reason is that many markets without bank branches in 1986 are calibrated as having low productivity, and these markets would have a much larger fraction of farmers than other markets if we directly used the calibrated subsistence income f for the alternative specification without productivity heterogeneity. Thus, to match the aggregate fraction of farmers in 1986, we need to calibrate a lower value of subsistence income f to make being a farmer less attractive.

Table OA.13: Correct prediction ratio in 1996, accounting for intramarket distances.

Specification	$d_{i,intra}$	Whole country	Bangkok	Central	North	Northeast	South
1	0 (baseline)	86.3%	87.5%	83.2%	83.0%	89.1%	83.6%
2	radius of a circle	84.7%	79.4%	84.0%	78.9%	89.8%	78.1%
3	diameter of a circle	80.7%	77.8%	80.3%	77.7%	81.9%	74.0%
4	side of a square	83.8%	82.4%	83.5%	80.9%	85.2%	79.5%

Note: This table presents the correct prediction ratio for branch locations in 1996 in different regions of Thailand. We calculate the distance from the nearest branch $d_{i,t}$ according to equation (4.1). Specification 1 corresponds to the predictions of our baseline model, where $d_{i,intra} \equiv 0$. We consider three alternative ways to measure $d_{i,intra}$ for each market i in other specifications. In specification 2, we assume that $d_{i,intra}$ is equal to the radius of a circle whose area is equal to the area of market i, i.e., $\pi d_{i,intra}^2 = Area_i$. In specification 3, we assume that $d_{i,intra}$ is equal to the diameter of the circle, i.e., $\pi(d_{i,intra}/2)^2 = Area_i$. In specification 4, we assume that $d_{i,intra}$ is equal to the length of one side of a square whose area is equal to the area of market i, i.e., $d_{i,intra}^2 = Area_i$.

subsistence income f, these markets generally already have a high fraction of workers (and a low fraction of farmers) relative to markets in the alternative model specification without productivity heterogeneity. As a result, the estimated impact of branch openings on local employment in our baseline model is smaller than that in the alternative specification without productivity heterogeneity (panel A of Figure OA.34). Once the smaller impacts of branch openings on local employment are aggregated across markets, our baseline model naturally implies that bank expansion leads to a less significant increase in the fraction of workers and a less significant decrease in the fraction of farmers compared with the alternative model specification without productivity heterogeneity (panels E and F of Figure OA.35). The weaker labor transition from farming into wage labor at the aggregate level further implies lower GDP growth over the 1986-1996 period compared with the model without productivity heterogeneity.

Figure OA.36 shows that our baseline model has higher overall income inequality than the alternative model without productivity heterogeneity. This is mainly due to the higher income differences across markets (see panel B of Figure OA.37) as a result of productivity heterogeneity. Panel A of Figure OA.37 shows that the within-market income inequality does not differ greatly between our baseline model and the alternative model specification without productivity heterogeneity.

4.3.2 Predictions of Branch Locations with Intramarket Distances

Our baseline model does not consider the costs of traveling within markets as the distance from the nearest branch $d_{i,t}$ is assumed to be 0 (equation (3.1) in the main text) in the case of markets where there is a local branch. In principle, this assumption means that there is more incentive to open branches in markets with larger areas in the data, and there are fewer such markets around

Bangkok and more in other regions. In this subsection, we evaluate the model's predictions of branch locations if intramarket distances are taken into account.

Specifically, we modify the calculation for the distance from the nearest branch $d_{i,t}$ defined in Section 3.1 of the main text as follows. For $i \in \Psi_t$, we have $d_{i,t} = d_{i,intra}$, and for $i \notin \Psi_t$, we have

$$d_{i,t} = \min \{\tau_{ij} : j \in \Psi_t, \Pi_{j,t} < h\} + d_{i,intra}, \tag{4.1}$$

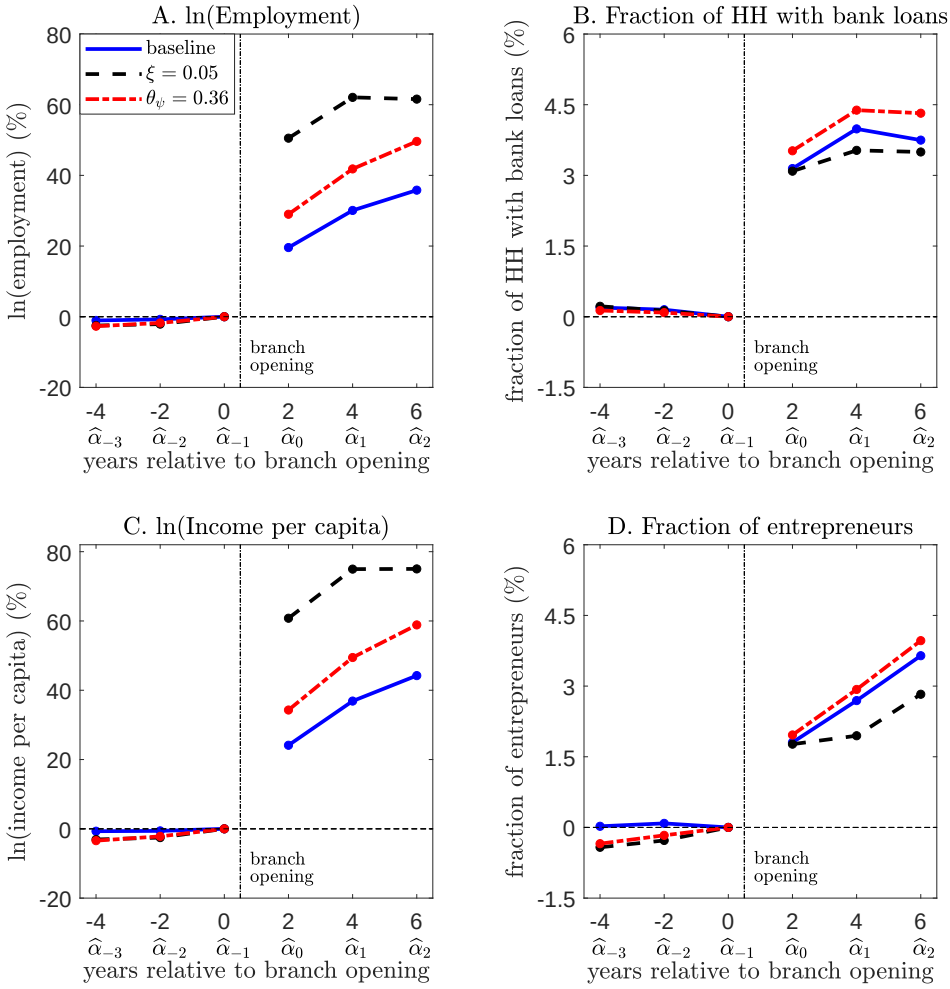
where $d_{i,intra}$ captures the travel distance within market i. If we assume that $d_{i,intra} \equiv 0$, the distance $d_{i,t}$ defined in equation (4.1) would be the same as that in our baseline model, which ignores intramarket distances.

We consider three ways to measure $d_{i,intra}$ for each market i as alternatives to the baseline model (which we refer to here as specification 1). First, we assume that $d_{i,intra}$ is equal to the radius of a circle whose area is equal to the area of market i, i.e., $\pi d_{i,intra}^2 = Area_i$ (specification 2). Second, we assume it is equal to the diameter of the circle, i.e., $\pi(d_{i,intra}/2)^2 = Area_i$ (specification 3), and third, we assume it is equal to the length of one side of a square whose area is equal to the area of market i, i.e., $d_{i,intra}^2 = Area_i$ (specification 4).

Table OA.13 presents the correct prediction ratio for branch locations in 1996 in various regions of Thailand. Compared with the baseline model's predictions (i.e., specification 1), the overall correct prediction ratio decreases slightly from 86.3% to 84.7% in specification 2, where intramarket distances are measured by the radius of a circle. The predictions for the central and northeast regions improve relative to the baseline model, but the predictions for other regions are worse. The correct prediction ratios are even lower than the baseline model in specifications 3 and 4, although all of them have overall correct prediction ratios of more than 80%. The results in Table OA.13 indicate that accounting for intramarket distances does affect the model's predictions of branch locations, but it is not a main driver compared with the three market characteristics that we highlight in Section 5.2 of the main text. Our baseline model's predictions of branch locations are robust after accounting for intramarket distances.

4.3.3 Collateral Constraints and Distance Sensitivity

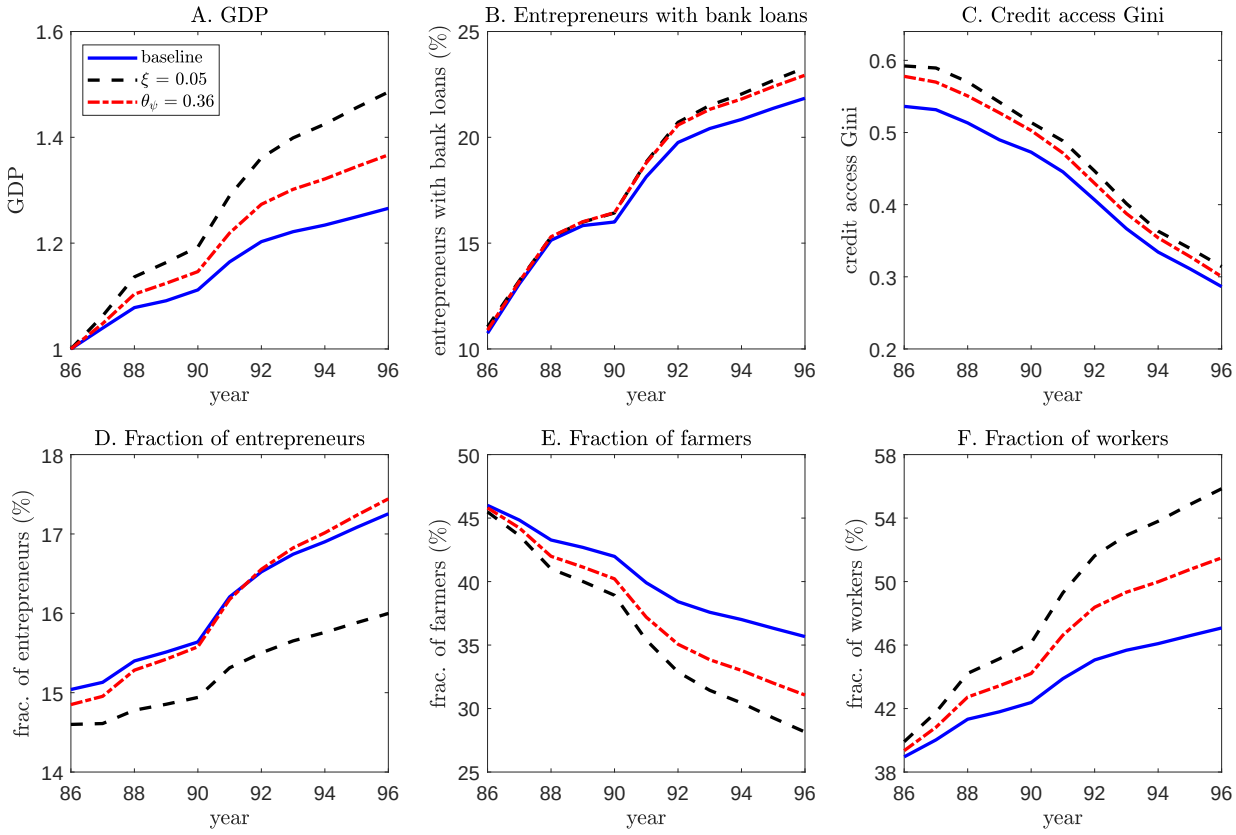
The parameters ξ and θ_ψ play important roles in determining the model-implied effects of bank expansion, both at the local market level and at the aggregate level. In our baseline calibration, these two parameters are calibrated at $\xi = 0.59$ and $\theta_\psi = 0.215$ so that the model-implied DID estimates for employment and credit access are consistent with those estimated in the data (see panel B of Table 2 of the main text). In this subsection, we study the sensitivity of the model's quantitative implications to each parameter.



Note: The blue solid lines represent the DID estimates in the baseline model (which are the same as the blue solid lines in Figure 5 of the main text). The black dashed lines represent the DID estimates under an alternative calibration with $\xi = 0.05$. Regarding the parameter θ_ψ, instead of recalibrating its value to match the DID estimate for credit access, we set $\theta_\psi = 0.215$ as in the baseline calibration in panel B of Table 2 of the main text for comparison purposes. The red dash-dotted lines represent the DID estimates under an alternative calibration with $\theta_\psi = 0.36$. Regarding the parameter ξ, instead of recalibrating its value to match the DID estimate for employment, we set $\xi = 0.59$ as in the baseline calibration in panel B of Table 2 of the main text for comparison purposes. For both the black dashed and red dash-dotted lines, we calibrate the other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, and θ_ζ) to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model.

Figure OA.38: Sensitivity to ξ and θ_ψ: Market-level effects of branch openings.

Role of ξ. To study the role of ξ, we consider an alternative calibration where we set $\xi = 0.05$ according to the structural estimate of Paulson, Townsend and Karaivanov (2006). For comparison with our baseline results, we set $\theta_\psi = 0.215$ at the same value as our baseline calibration. The other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline calibration. The black dashed



Note: The blue solid lines represent the aggregate dynamics in the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics under an alternative calibration with $\xi = 0.05$. The red dash-dotted lines represent the aggregate dynamics under an alternative calibration with $\theta_\psi = 0.36$. Other parameters are calibrated according to the note for Figure OA.38.

Figure OA.39: Sensitivity to ξ and θ_ψ: Aggregate dynamics, 1986-1996.

lines in Figure OA.38 plot the DID estimates from the simulated data based on the calibration with $\xi = 0.05$. Compared with our baseline calibration of $\xi = 0.59$ (blue solid lines), the impacts of branch openings on local employment (panel A) and income per capita (panel C) more than double under the calibration of $\xi = 0.05$, whereas the impacts on the fraction of households with access to credit (panel B) and the fraction of entrepreneurs (panel D) are almost unchanged. In Figure OA.39, we present the aggregate dynamics for the 1986-1996 period. Consistent with the larger market-level effects, under the calibration of $\xi = 0.05$, GDP grows by 48.7% from 1986 to 1996, which is about twice the GDP growth implied by our baseline calibration. Moreover, the model predicts that the fraction of workers will increase to 55.8% in 1996 if $\xi = 0.05$, compared with 47.0% under our baseline calibration of $\xi = 0.59$.

Intuitively, compared with our baseline calibration of $\xi = 0.59$, setting $\xi = 0.05$ implies that entrepreneurs will face much more relaxed borrowing constraints once they obtain access to credit.

This allows them to significantly expand their businesses and quickly accumulate wealth, resulting in large increases in their output and their demand for workers. This explains why new branch openings have more significant impacts on local output and employment under the calibration of $\xi = 0.05$ than under our baseline calibration of $\xi = 0.59$. Once these larger local impacts are aggregated across markets, bank expansion implies much more significant GDP growth at the aggregate level than under the baseline calibration. Furthermore, we find that the higher GDP growth over the 1986-1996 period implied by the calibration of $\xi = 0.05$ is caused by fast transitions rather than by high steady-state values. This is consistent with the insight of Buera and Shin (2013), who argue that tight financial constraints result in slow economic transitions in a calibrated quantitative model. However, in the long-run steady state, the output loss caused by tight financial constraints can be small because entrepreneurs optimally choose to accumulate wealth to self finance their businesses (Buera and Shin, 2011; Moll, 2014).

Role of θ_ψ. To study the role of θ_ψ, we consider an alternative calibration where we set a larger $\theta_\psi = 0.36$ (compared with $\theta_\psi = 0.215$ in the baseline calibration). For comparison with our baseline results, we set $\xi = 0.59$ at the same value as our baseline calibration. The other parameters (i.e., β, f, κ, η, ψ, ζ, ρ, γ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline calibration. The red dash-dotted lines in Figure OA.38 plot the DID estimates from the simulated data based on the calibration of $\theta_\psi = 0.36$. Compared with our baseline calibration of $\theta_\psi = 0.215$ (blue solid lines), calibrating a larger value of θ_ψ implies larger impacts of branch openings on all four outcome variables of interest because it makes the costs of accessing bank loans more sensitive to the distance from the nearest branch $d_{i,t}$. Importantly, while the parameter ξ has little impact on the fraction of households with bank loans (panel B) or the fraction of entrepreneurs (panel D), these two DID estimates, which reflect the impacts of branch openings on the extensive margin, are sensitive to the choice of θ_ψ. In fact, in our baseline calibration, the values of $\xi = 0.59$ and $\theta_\psi = 0.215$ are calibrated so that the model-implied DID estimates for employment and credit access are consistent with those estimated in the data (see panel B of Table 2 of the main text).

Because a larger θ_ψ implies that branch openings will have more significant local impacts, we expect that the model-predicted transitional dynamics will exhibit larger variations over time than the baseline calibration. For example, the red dash-dotted line in panel A of Figure OA.39 shows that GDP will grow by 36.7% between 1986 and 1996 under the calibration with $\theta_\psi = 0.36$ compared with the 26.5% GDP growth under our baseline calibration of $\theta_\psi = 0.215$ (blue solid line).

In a complementary analysis to the sensitivity analyses of θ_ψ, in Online Appendix 4.2.2, we show that the value of θ_ζ, which governs the sensitivity of portfolio adjustment costs to $d_{i,t}$, does not have significant effects on the model-implied local and aggregate effects of bank expansion

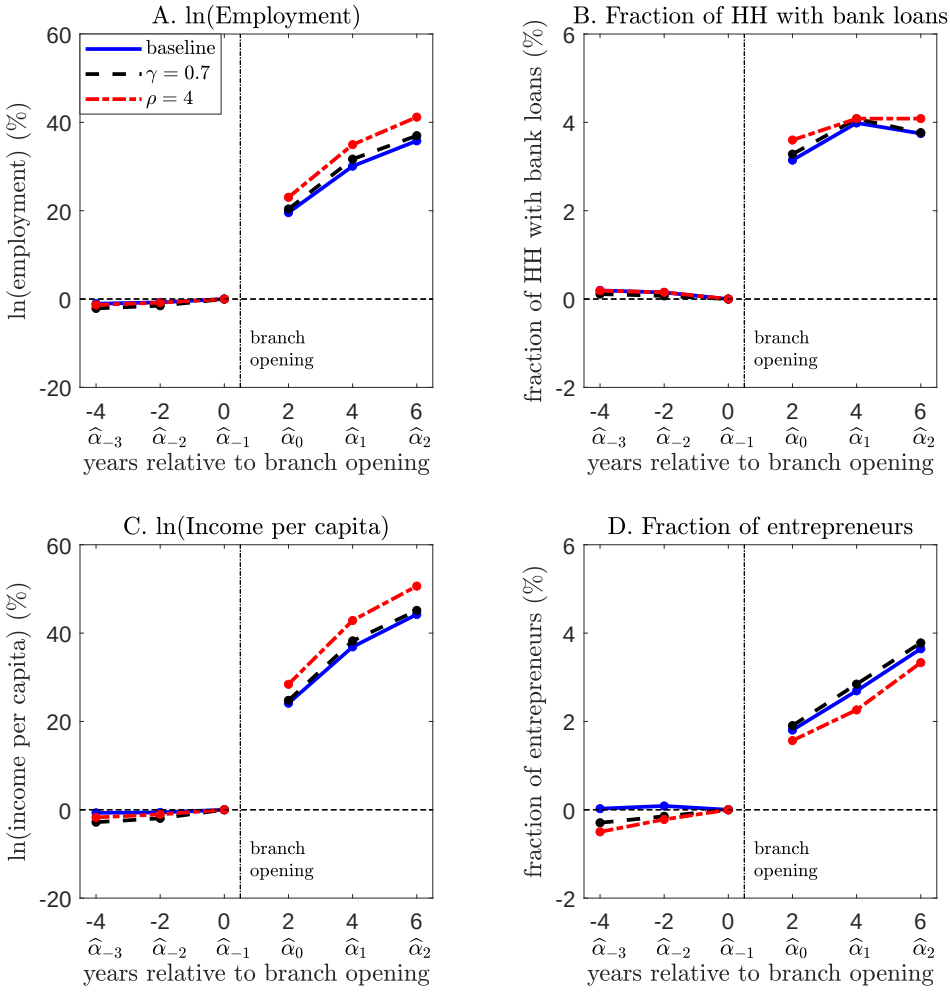
once the two key parameters, θ_ψ and ξ, are recalibrated to match the DID estimates for the 2-year effect of branch openings on local credit access and employment.

4.3.4 Stochastic Process of Talent

We study the sensitivity of the model's quantitative implications to parameters γ and ρ. These two parameters determine the persistence of and variation in individual talent, respectively, which determine the revenue generated by household assets. In our baseline calibration, these parameters are calibrated at $\gamma = 0.72$ and $\rho = 4.7$ so that the overall dynamics of households' asset-generated revenue implied by the model are consistent with those in the Townsend Thai annual survey. We choose four moments to capture the dynamics of asset-generated revenue, namely the standard deviation of log revenue growth and the autocorrelation of log revenue for horizons of 1, 3, and 5 years.

We can calibrate the parameters γ and ρ in alternative ways. Intuitively, the persistence of talent γ determines the frequency with which entrepreneurs quit their businesses. The Townsend Thai annual survey indicates that entrepreneurs remain in their current businesses for about 4 years. Thus, as an alternative calibration, we set $\gamma = 0.7$ so that the model-implied moment is about 4 years, as in the data. The tail parameter ρ can be calibrated based on the cross-sectional size distribution. Based on the World Bank Enterprise Survey, we estimate that the top 20% of Thai manufacturing firms have a 0.72 employment share. Thus, an alternative calibration for ρ is to set $\rho = 4$ to generate roughly the same employment share in our model. In the rest of this section, we study the implications of these alternative calibrations. We show that under the alternative calibration, the model-implied aggregate impacts of bank expansion on GDP growth are larger than those under our baseline calibration of $\gamma = 0.72$ and $\rho = 4.7$, but the difference is not sizable. Thus, our model implications are robust to alternative calibrations of these two parameters.

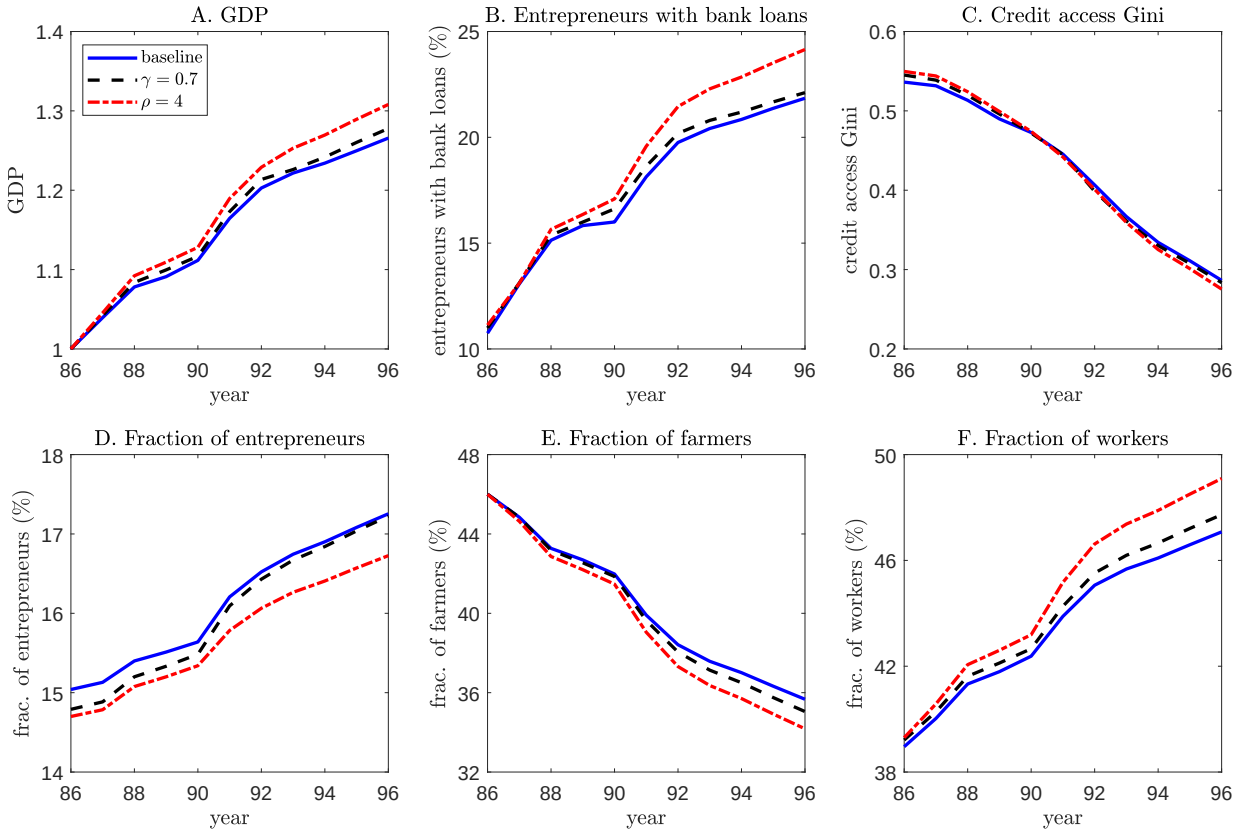
Role of γ. To study the role of γ, we consider an alternative calibration where we set $\gamma = 0.7$ to capture a less persistent talent process than under the baseline calibration, and retain the same value for ρ as our baseline calibration, i.e., $\rho = 4.7$. Because the two parameters θ_ψ and ξ crucially determine the quantitative impacts of bank expansion, for comparison with our baseline model, we set $\theta_\psi = 0.215$ and $\xi = 0.59$ at the same values as in our baseline calibration. The other parameters (i.e., β, f, κ, η, ψ, ζ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 in the main text, as in the baseline calibration. The black dashed lines in Figure OA.40 plot the DID estimates from the simulated data based on the calibration with $\gamma = 0.7$. Compared with our baseline calibration with $\gamma = 0.72$ (blue solid lines), the impacts of branch openings on local employment (panel A) and income per capita (panel C) are slightly



Note: The blue solid lines represent the DID estimates in the baseline model (which are the same as the blue solid lines in Figure 5 of the main text). The black dashed lines represent the DID estimates under an alternative calibration with $\gamma = 0.7$, and we set $\rho = 4.7$ at the same value as in our baseline calibration. The red dash-dotted lines represent the DID estimates under an alternative calibration with $\rho = 4$, and we set $\gamma = 0.72$ at the same value as our baseline calibration. For both the black dashed and red dash-dotted lines, instead of recalibrating the two parameters θ_ψ and ξ to match the DID estimates for credit access and employment, we set $\theta_\psi = 0.215$ and $\xi = 0.59$ at the same values as in our baseline calibration for comparison purposes. The other parameters (i.e., β, f, κ, η, ψ, ζ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 of the main text, as in the baseline model.

Figure OA.40: Sensitivity to γ and ρ: Market-level effects of branch openings.

larger under the calibration with $\gamma = 0.7$. In Figure OA.41, we present the aggregate dynamics for the 1986-1996 period. Consistent with the larger market-level effects on income per capita, under the calibration with $\gamma = 0.7$, GDP would grow by 27.7% over 1986-1996 (the black dashed line in panel A), whereas under our baseline calibration (the blue solid line in panel A), GDP growth is about 26.5%. Moreover, the fraction of entrepreneurs in 1986 is 14.8% under the calibration with $\gamma = 0.7$ (the black dashed line in panel D), smaller than the 15.0% implied by our baseline



Note: The blue solid lines represent the aggregate dynamics in the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics under an alternative calibration with $\gamma = 0.7$. The red dash-dotted lines represent the aggregate dynamics under an alternative calibration with $\rho = 4$. Other parameters are calibrated according to the note for Figure OA.40.

Figure OA.41: Sensitivity to γ and ρ: Aggregate dynamics, 1986-1996.

calibration (the blue solid line in panel D). In both calibrations, we choose the parameter ψ to match the fraction of households with bank loans in 1986 in the data (1.63%). Thus, a lower fraction of entrepreneurs in 1986 implies a slightly higher fraction of entrepreneurs with bank loans under the calibration with $\gamma = 0.7$ (the black dashed line in panel B).

Intuitively, under the calibration with $\gamma = 0.7$, households find it more difficult to overcome their borrowing constraints because they have less persistent talent process compared with the baseline calibration with $\gamma = 0.72$. Thus, in 1986, fewer households choose to be entrepreneurs and more choose to be workers compared with the baseline calibration. The economy's level of GDP in 1986 is lower than the baseline calibration with $\gamma = 0.72$, consistent with the insight of Buera and Shin (2011) and Moll (2014) that financial frictions result in larger total factor productivity and output losses in the steady state when firm productivity is less persistent. The lower initial level of GDP implies a higher GDP growth rate in subsequent years as bank expansion

gradually alleviates financial frictions across regions over time.

Role of ρ. To study the role of ρ, we consider an alternative calibration where we set the tail parameter at $\rho = 4$ to capture a more dispersed talent distribution, and retain $\gamma = 0.72$ at the same value as in our baseline calibration. Because the two parameters θ_ψ and ξ crucially determine the quantitative impacts of bank expansion, for comparison with our baseline model, we set $\theta_\psi = 0.215$ and $\xi = 0.59$ at the same values as in our baseline calibration. The other parameters (i.e., β, f, κ, η, ψ, ζ, and θ_ζ) are calibrated to match the corresponding moments in panel B of Table 2 in the main text, as in the baseline calibration. The red dash-dotted lines in Figure OA.40 plot the DID estimates from the simulated data based on the calibration with $\rho = 4$. Compared with our baseline calibration with $\rho = 4.7$ (blue solid lines), under the calibration with $\rho = 4$, the impacts of branch openings on local employment (panel A), credit access (panel B), and income per capita (panel C) are larger whereas the impacts on the local fraction of entrepreneurs are smaller. In Figure OA.41, we present the aggregate dynamics over the 1986-1996 period. Consistent with the larger market-level effects on income per capita, under the calibration with $\rho = 4$, GDP would grow by 30.0% between 1986 and 1996 (red dash-dotted line in panel A) compared with 26.5% under our baseline calibration (blue solid line in panel A). Moreover, the fraction of entrepreneurs with bank loans (red dash-dotted line in panel B) and the fraction of workers (red dash-dotted line in panel F) also increase more dramatically under the calibration of $\rho = 4$.

Intuitively, calibrating a smaller value of $\rho = 4$ makes the Pareto distribution of individual talent more right skewed than under the baseline calibration, so the fraction of households with high talent levels increases. These high-talent households operate large-scale businesses and drive up the equilibrium local wages in every market. The higher wages motivate households with mediocre talent to be workers rather than entrepreneurs, resulting in a lower fraction of entrepreneurs and a higher fraction of workers in 1986 (as shown by the red dash-dotted lines in panels D and F, respectively, of Figure OA.41). In addition, the high-talent households are more constrained by any collateral constraints because they face a high marginal return of capital in this scenario. As a result, when financial frictions are alleviated by bank expansion, these high-talent households will choose to borrow from banks to significantly expand their productive capital and hire more workers, leading to larger increases in aggregate GDP, the fraction of workers, and the fraction of entrepreneurs with bank loans. However, the impacts of branch openings on the fraction of entrepreneurs are less significant under the calibration of $\rho = 4$ due to the dramatic increase in local wages.

Table OA.14: Net capital flows to Thailand

	86	87	88	89	90	91	92	93	94	95
ς_t (%)	-0.80	2.20	6.10	8.20	13.00	10.70	8.40	8.40	8.60	13.10

Note: This table presents the net capital flows as a percent of GDP between 1986 and 1995 in Thailand. The data are from the Bank of Thailand (Alba, Hernandez and Klingebiel, 1999, Figure 2), which includes nonbank and bank capital flows. The nonbank capital flows include foreign direct investment (FDI), portfolio capital, nonresident baht accounts, trade credits, and syndicated borrowing by domestic corporates from overseas financial institutions. The bank capital flows represent resident banks borrowing from overseas sources.

4.3.5 Extension with International Capital Flows

Between 1988 and 1996, Thailand received very large capital inflows relative to its GDP. As summarized in Table OA.14, there were no net capital flows in 1986, but they surged from 1988, and by 1990 had reached 13% of GDP. A further small rise occurred in 1995, with capital flows reaching 13.1% of GDP.

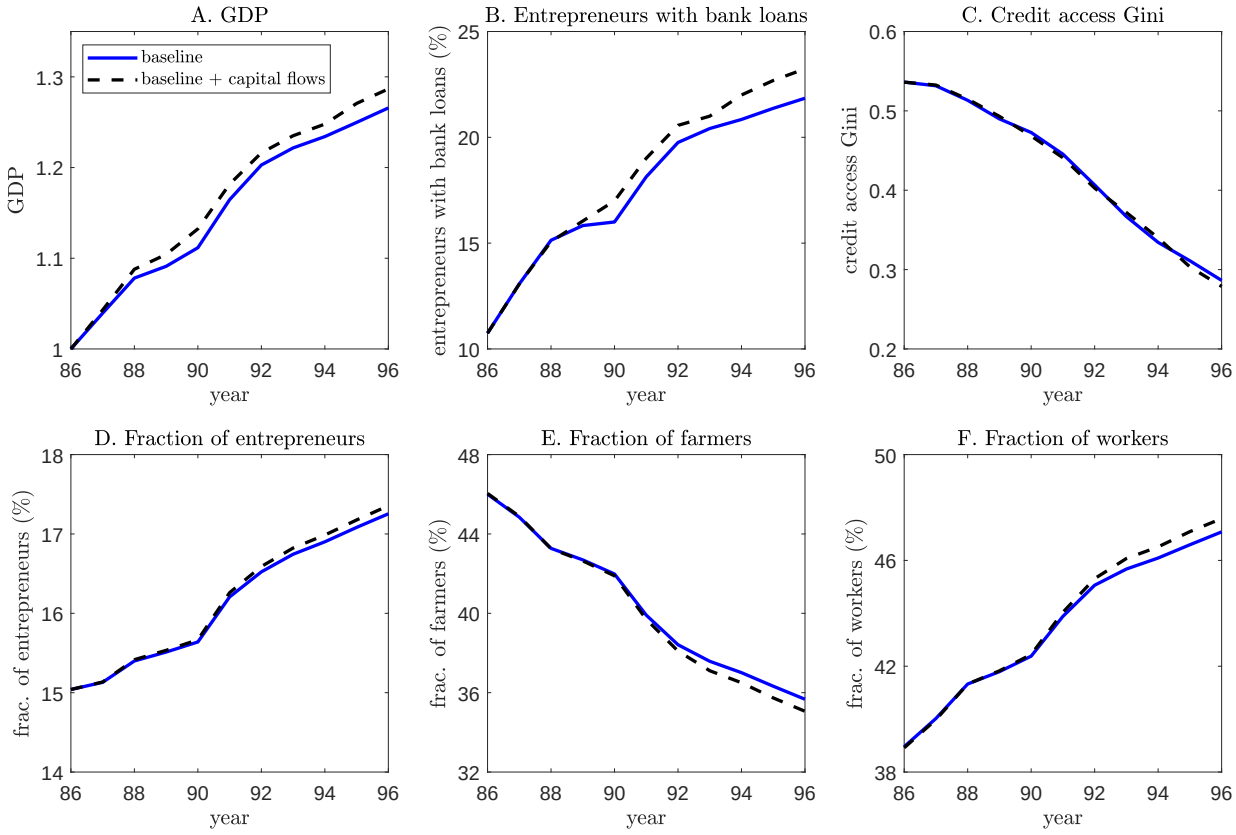
The surge in net capital flows may have had a direct positive impact on GDP growth and fueled the boom in credit access during our study period. To evaluate this, we extend the baseline model to incorporate the effect of international capital flows on the supply of capital. Specifically, the capital-market clearing condition in equation (3.18) of the main text is modified as follows:

$$\sum_{i=1}^{N} \Pi_{i,t} \int a\phi_t(s|i)\mathrm{d}s + \varsigma_t Y_t = \sum_{i=1}^{N} X_{i,t}, \tag{4.2}$$

where ς_t is net capital flows as a percentage of GDP and Y_t is GDP in year t, defined in equation (3.28) in Online Appendix 3.6.1. An increase in ς_t, which captures the surge in net capital flows to Thailand, lowers the equilibrium interest rate r_t by increasing the supply of capital.

In Figure OA.42, the blue solid lines plot the aggregate implications of our baseline model, which does not account for capital flows, and the black dashed lines plot the aggregate implications of the extended model with capital flows. Comparing the blue solid and black dashed lines in panel A, GDP growth for the 1986-1996 period increases from 26.5% in our baseline model to 28.6% in the extended model with capital flows. Thus, net capital flows result in GDP growth of around 2.1%. We find that the effect of capital flows in our extended model is small because capital flows influence the real economy only by increasing the supply of capital through equation (4.2). Our extended model does not capture many other important channels through which capital flows may promote economic growth. For example, capital flows, particularly FDI, may contribute to productivity growth through technology transfers. Moreover, capital account liberalization may influence internal policy reforms.

Panel B of Figure OA.42 shows that by increasing the supply of capital and lowering the equilibrium interest rate, the surge in net capital flows increases the fraction of entrepreneurs



Note: The blue solid lines represent the aggregate dynamics of the baseline model (which are the same as the blue solid lines in Figure 8 of the main text). The black dashed lines represent the aggregate dynamics under an alternative calibration with international capital flows, which results in an increase in the aggregate supply of capital. Capital flows to Thailand are calibrated according to Table OA.14.

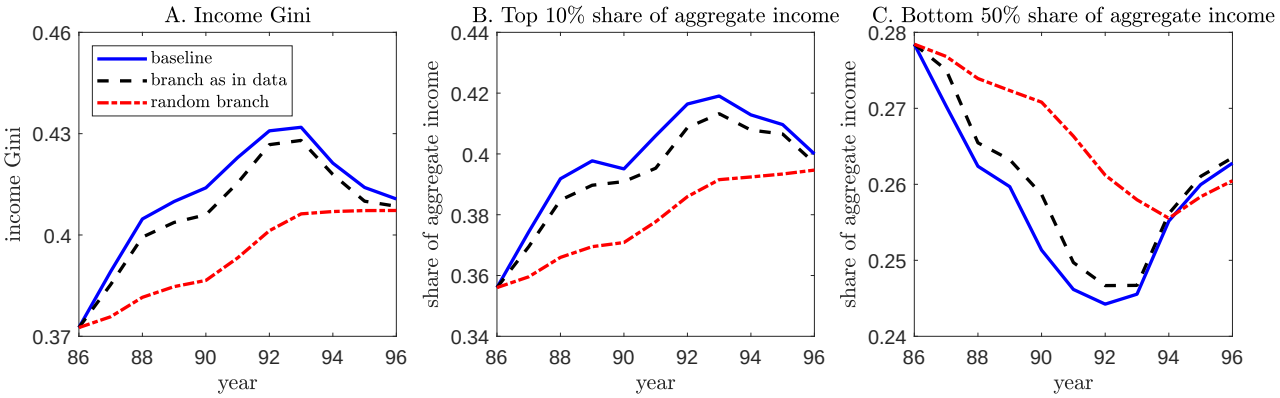
Figure OA.42: International capital flows: Aggregate dynamics, 1986-1996.

with access to bank loans. However, panel C of Figure OA.42 shows that credit access inequality across markets is virtually unaffected by capital flows. Panels D to F of Figure OA.42 show that the surge in capital flows leads to a higher fraction of entrepreneurs and workers, and a lower fraction of farmers in the economy.

4.4 Exogenous Branch Locations

In this section of the appendix, we show that the model's quantitative implications are similar if branch locations are exogenously given by the data, because our baseline model already correctly predicts the majority of branch locations.

In particular, we solve an alternative model specification in which new branches opened between 1987 and 1996 are exogenously placed in the same locations as those in the data. That is, we solve households' optimal decisions taking the branch locations from the data in each year



Note: The blue solid lines represent the dynamics in the baseline model (which are the same as the blue solid lines in Figure 9 of the main text). The black dashed lines represent the dynamics in an alternative model specification in which new branches opened between 1987 and 1996 are exogenously placed at the same locations as those in the data. The red dash-dotted lines represent the dynamics in an alternative model specification in which new branch locations are randomly selected.

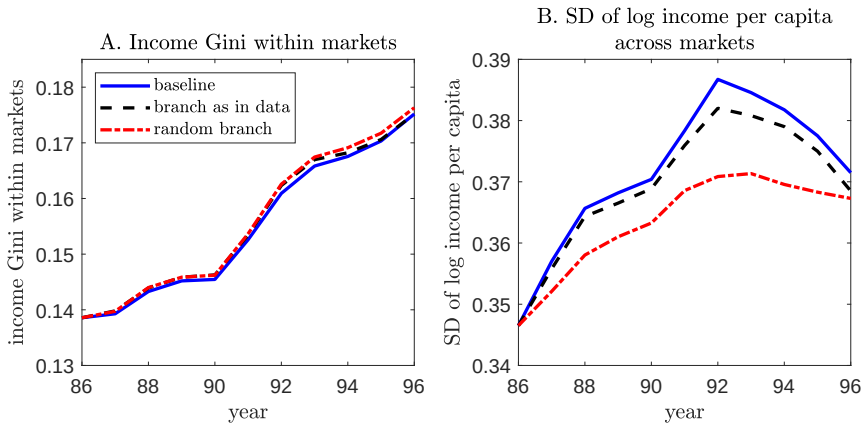
Figure OA.43: Exogenous and random branch locations: Overall income inequality across households.

as given. Figures OA.43 and OA.44 compare the income inequality dynamics predicted by our baseline model (the blue solid lines) and this alternative model specification (the red dash-dotted lines).

In terms of the overall income inequality dynamics (panel A of Figure OA.43) and the income inequality dynamics across markets (panel B of Figure OA.44), the black dashed lines display an inverted U-shape similar to that of the blue solid lines. However, the peak levels of the overall and cross-market income inequality are slightly higher in our baseline specification than in the alternative specification, because in the baseline specification, branches tend to be opened in high-productivity markets (see the discussions in Online Appendix 3.7.2). Households in high-productivity markets already earn higher income than households in other markets and thus opening more branches in such markets will lead to higher income inequality across markets. In terms of the inequality dynamics within markets (panel A of Figure OA.44), the black dashed line and blue solid line are almost overlapping because the locations of branch locations have little effect on within-market income differences across households.

Panel A of Figure OA.45 presents the cross-market welfare gains implied by this alternative model specification. Overall country-wide welfare increases by 17.8% from 1986 to 1996, compared with 19.9% implied by our baseline model (Section 6.3 of the main text).

Figure OA.46 presents the aggregate dynamics predicted by this alternative model specification. The black dashed line shows that GDP growth between 1986 and 1996 would be slightly lower than that predicted by our baseline model because branch locations are not chosen to maximize



Note: The blue solid lines represent the dynamics in the baseline model (which are the same as the blue solid lines in Figure 10 of the main text). The black dashed lines represent the dynamics in an alternative model specification in which new branches opened between 1987 and 1996 are exogenously placed at the same locations as those in the data. The red dash-dotted lines represent the dynamics in an alternative model specification in which new branch locations are randomly selected.

Figure OA.44: Exogenous and random branch locations: Income inequality within and across markets.

total profits in equation (3.14).[24] For the same reason, the changes in other aggregate variables (panels B to F) predicted by this alternative model specification are less pronounced compared with the baseline model's predictions. However, the difference is small because our baseline model correctly predicts the locations of 86.3% branches opened between 1987 and 1996.

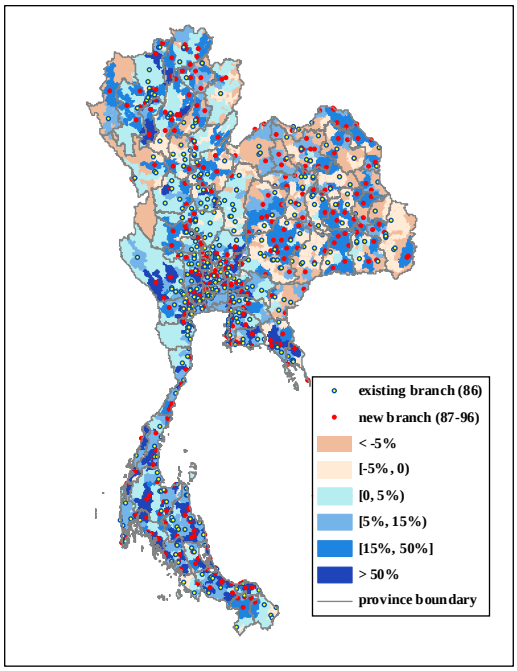
4.5 Randomly Selected Branch Locations

In this section of the appendix, we present the predictions of an alternative model specification in which the locations of new branches opened between 1987 and 1996 are randomly selected. In terms of branch locations, we show that our baseline model significantly outperforms this alternative model specification, which does not have any predictive power. In terms of quantitative implications, we show that this alternative model specification is significantly worse than our baseline model in fitting the data. This highlights the importance of predicting branch locations based on variations in market characteristics.

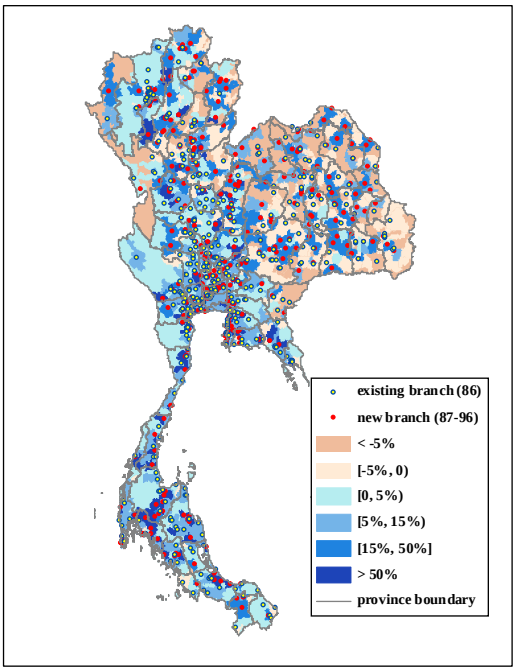
Specifically, we randomly select the locations of 431 new branches opened between 1987 and 1996 from among the 1,022 markets without branches in 1986. Panel A of Figure OA.47 plots the distribution of the correct prediction ratio over $1,000,000$ independent simulations. The median correct prediction ratio is 42.2%, with a very tight 95% confidence interval of $[38.8\%, 45.7\%]$. None of the simulations can generate a correct prediction ratio higher than that of our baseline

[24] In Figure OA.12 in Online Appendix 3.7.2, we show that opening branches in markets with the largest increase in loan demand (i.e., maximize bank profits) will also generally lead to the largest increase in local output.

A. Branch locations as in data



B. Random branch locations

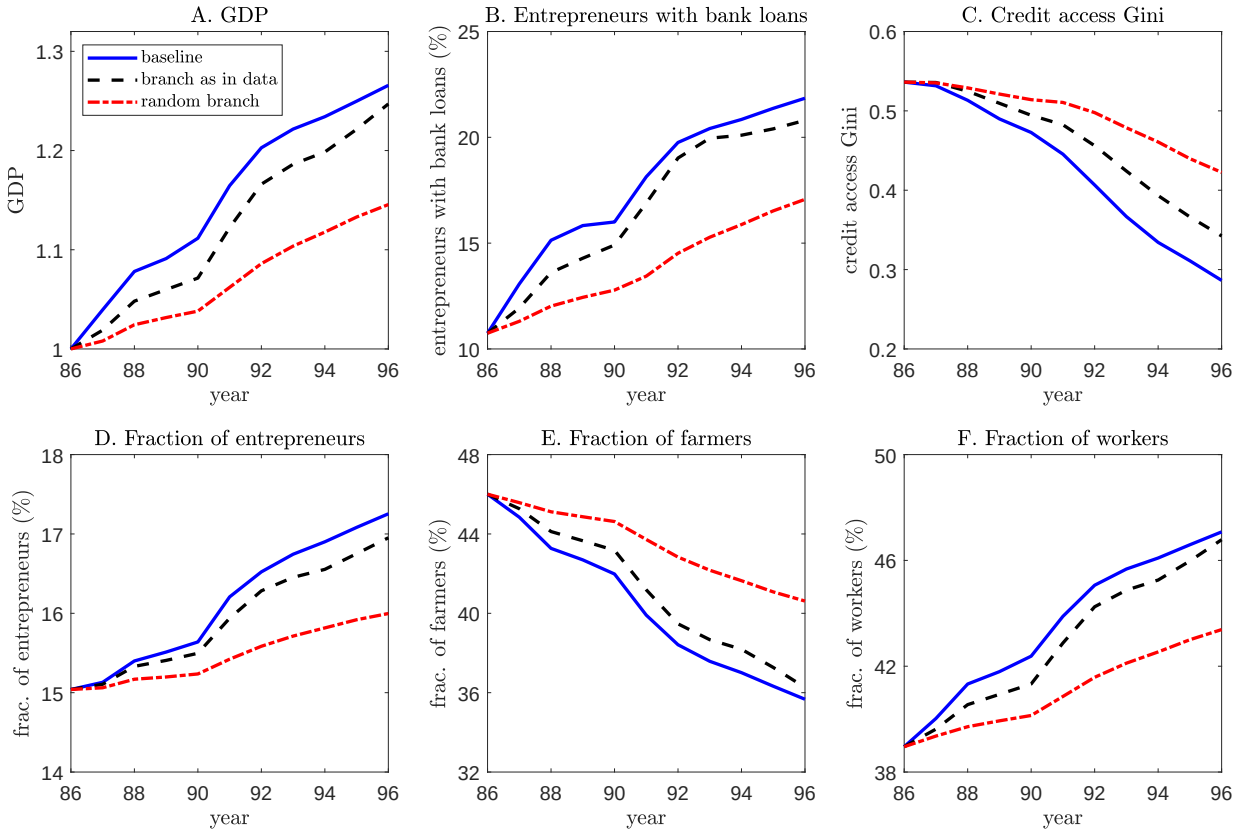


Note: This figure complements panel A of Figure 12 of the main text by presenting the cross-market welfare gains implied by two alternative model specifications. Panel A plots welfare gains in an alternative model specification in which new branches opened between 1987 and 1996 are exogenously placed at the same locations as those in the data. Overall country-wide welfare increases by 17.8% from 1986 to 1996. Panel B plots the welfare gains in an alternative model specification where new branch locations are randomly selected. Overall country-wide welfare increases by 10.9% from 1986 to 1996.

Figure OA.45: Model-implied market-level welfare gains from 1986 to 1996 based on alternative model specifications.

model (86.3%). Panel B plots the distribution of timing difference over $1,000,000$ independent simulations. The median timing difference is 7.05 years, with a very tight 95% confidence interval of $[6.76, 7.32]$. By contrast, the average timing difference is 2.38 years in our baseline model.

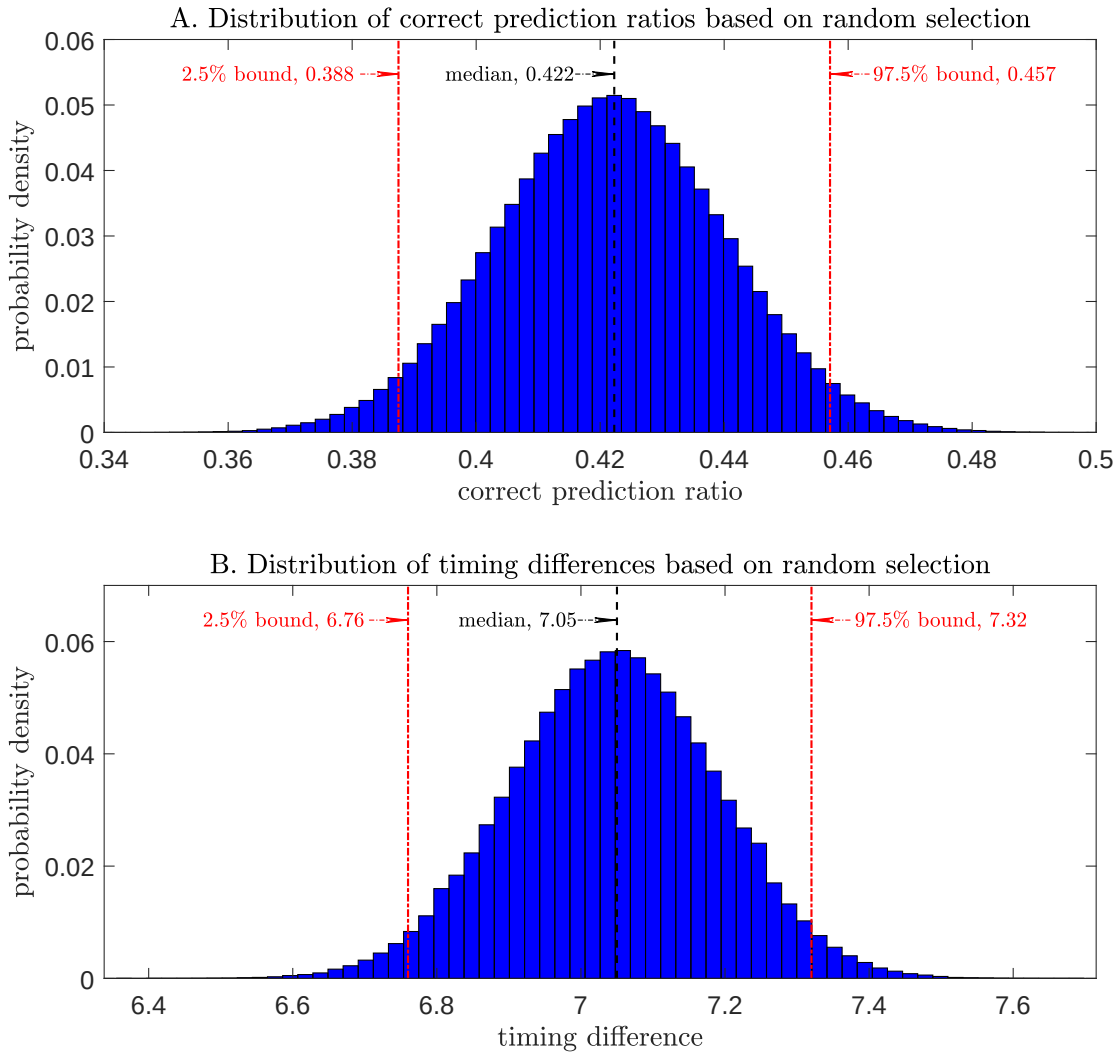
Next, we evaluate the quantitative implications of this alternative model specification. We focus on the allocation that delivers the median correct prediction ratio over 1,000,000 independent simulations (see panel A of Figure OA.47). Figures OA.43 and OA.44 compare the income inequality dynamics predicted by our baseline model (blue solid lines) and this alternative specification (red dash-dotted lines). The red dash-dotted line in panel A of Figure OA.43 shows that in this alternative model specification, the overall income Gini coefficient increases almost monotonically from 1986 to 1996, exhibiting a pattern that is very different from the inverted U-shape in our baseline model and the data. The main reason for this pattern, as indicated by the red dash-dotted line in panel B of Figure OA.44, is that the increase in the cross-market income Gini coefficient during 1986-1991 is smaller than in our baseline model (the blue solid



Note: The blue solid lines represent our baseline model; the black dashed lines represent an alternative model specification in which new branches opened between 1987 and 1996 are exogenously placed at the same locations as those in the data; and the red dash-dotted lines represent the median prediction of 1,000,000 independent simulations based on an alternative model specification in which new branch locations are randomly selected.

Figure OA.46: Aggregate dynamics, 1986-1996.

line). As a result, the cross-market income inequality dynamics display a less significant inverted U-shape in this alternative model specification. This is because in our baseline model, the central authority tends to open branches in high-productivity markets (see the discussions in Online Appendix 3.7.2). Households in such markets already earn higher income than households in other markets and thus opening more branches in such markets will lead to higher income inequality across markets. This increases the peak level of the cross-market income Gini coefficient, yielding a stronger inverted U-shape in the cross-market income inequality dynamics. Conversely, the within-market income inequality dynamics are similar in the baseline model and this alternative model specification (Figure OA.44) because branch locations have little effect on within-market income differences across households. The overall income Gini coefficient in panel A of Figure OA.43 is determined by both cross-market income inequality and within market income inequality. Thus, when the cross-market income inequality dynamics display a less significant inverted



Note: We randomly choose the locations of 431 new branches opened from 1987 to 1996 from among the 1,022 markets without branches in 1986. Panel A plots the distribution of the correct prediction ratio over $1,000,000$ independent simulations. The correct prediction ratio is defined as the number of model-predicted branch locations that are consistent with the data in 1996 divided by 431. Panel B plots the distribution of timing difference over $1,000,000$ independent simulations. For each new branch opened between 1987 and 1996 in the data, we compute the difference between its opening year and the model-predicted branch opening year. We set this difference to 10 years for branches that were opened in the data by 1996 but are not predicted by the model.

Figure OA.47: The distribution of correct prediction ratios and timing differences when branch locations are randomly selected.

U-shape in this alternative model specification compared with the baseline model, the overall income Gini coefficient will monotonically increase from 1986 and 1996, reflecting the strong monotonically increasing within-market income Gini coefficient (panel A of Figure OA.44).

Panel B of Figure OA.45 presents the cross-market welfare gains implied by this alternative model specification. Overall country-wide welfare increases by 10.9% from 1986 to 1996, compared with the 19.9% implied by our baseline model (Section 6.3 of the main text). In Figure OA.46, the red dash-dotted lines plot the aggregate dynamics implied by this alternative model specification. The GDP growth between 1986 and 1996 is about half of that predicted by the baseline model. The composition of occupations also changes much less dramatically compared with that in our baseline model and the data.

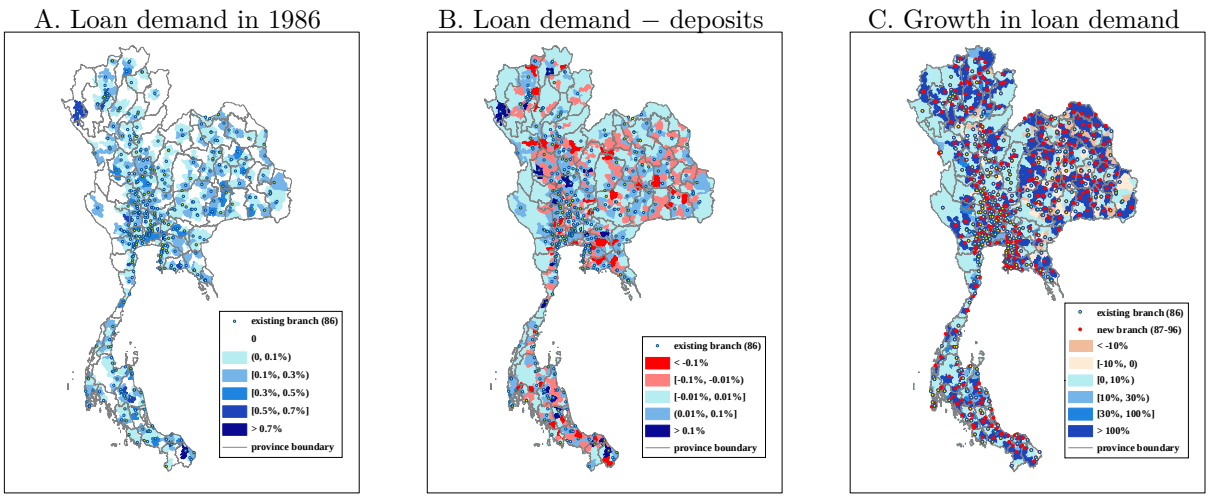
4.6 Flow of Funds Across Markets

In this section of the appendix, we present the flow of funds across markets implied by our model. Panel A of Figure OA.48 shows that most markets in 1986 lack access to bank loans (the white area of the figure) because they are far from existing branches. As a result, credit is also very unequally distributed across markets. Such regional heterogeneity underlies the low credit access ratio and the high credit access Gini coefficient for the aggregate economy (see panels D and E, respectively, of Figure 8 in the main text).

Because our model assumes free capital flows across markets, the local demand for bank loans is not necessarily equal to the local supply of funds (i.e., deposits). Panel B of Figure OA.48 shows that the local demand for loans is higher than the local amounts of deposits in regions close to bank branches (dark/light blue), indicating that capital flows to regions with better access to finance. The endogenous flow of funds implied by our model captures an important aspect of the within-country inter-regional capital flows in Thailand (Paweenawat and Townsend, 2019), as it contributes to high inequality among entrepreneurs in different regions. The rapid expansion of banks during the study period leads to a significantly higher density of bank branches in 1996. In panel C of Figure OA.48, the red dots represent bank branches newly opened between 1987 and 1996 predicted by our model. The markets close to these branches experience a dramatic increase in local loan demand, as represented by the dark blue color.

4.7 Alternative Specification with Bank Expansion until 2011

In our baseline specification (Figure 8 of the main text), we assume that the distribution of bank branches is fixed after 1996, and that no more new branches will be opened after 1997. In this appendix section, we show that this assumption has a negligible impact on the transitional dynamics between 1986 and 1996, our period of interest. In Online Appendix 4.8, we deal



Note: Panel A plots the percentage of bank loans (i.e., local demand for loans divided by the aggregate amount of bank loans) borrowed by entrepreneurs in each local market in 1986. Panel B plots the local excess demand for loans in 1986, which is defined as the difference between the local demand for loans and the local supply of funds (i.e., deposits) divided by the aggregate amount of bank deposits. The red regions indicate excess local supply of funds, and the blue regions indicate excess local demand for loans. Panel C plots the growth in local loan demand from 1986 to 1996.

Figure OA.48: Model-implied loan demand and deposits.

explicitly with the post 1996 financial crisis.

Figure OA.49 plots the numbers of branches opened and closed in each year in the data from 1987 to 2011. The black dashed line shows that for the 1986-1996 period, there are no branch closures, consistent with the baseline model. However, after 1999, a number of branches were closed. Rysman, Townsend and Walsh (2022) attribute the branch closures to the financial crisis.[25] Due to the large number of branch closures and the small number of new branches opened between 1999 and 2002, the net change in the number of branches is negative in these years (the red dash-dotted line).

To formally establish the negligible impact of branches opened after 1996 on the transitional dynamics in the 1986-1996 period, we solve the model with an alternative specification in which new branches continue to open from 1987 to 2011, the last year for which data are available to us. We follow the baseline calibration and take the number of new branches in each year from 1997 to 2011 as the same as the net change in the number of branches in the data (the red dash-dotted line in Figure OA.49). Because our model does not allow for branch closures, we calibrate the number of new branches between 1999 and 2002 to be zero. The absence of branch closures does not affect the quantitative implications of the model much for the 1986-1996 period because

[25] Their structural model implies that there would have been 15% more branches and 9% more markets with at least one branch after 10 years had the crisis not occurred.

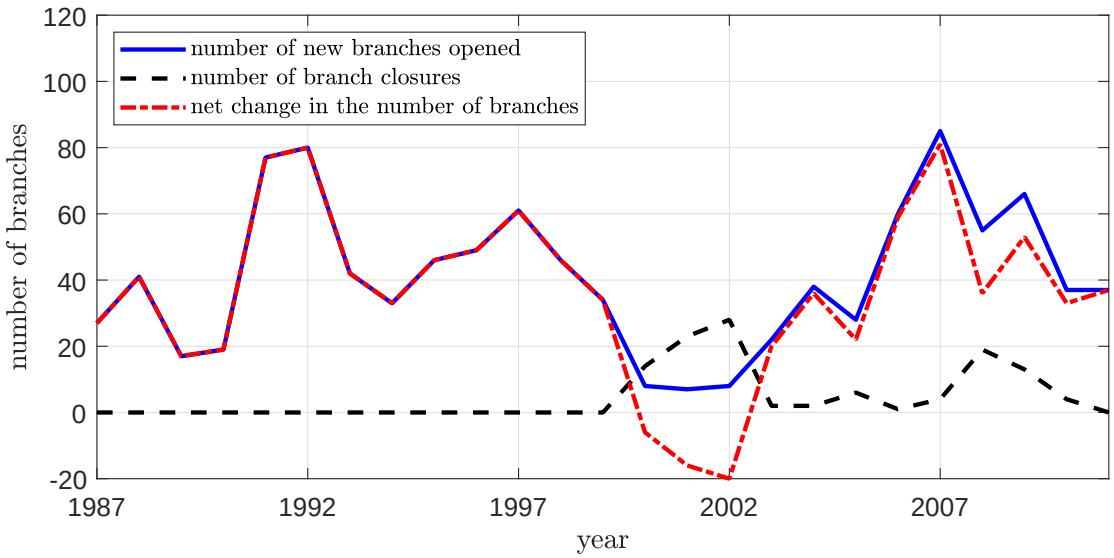
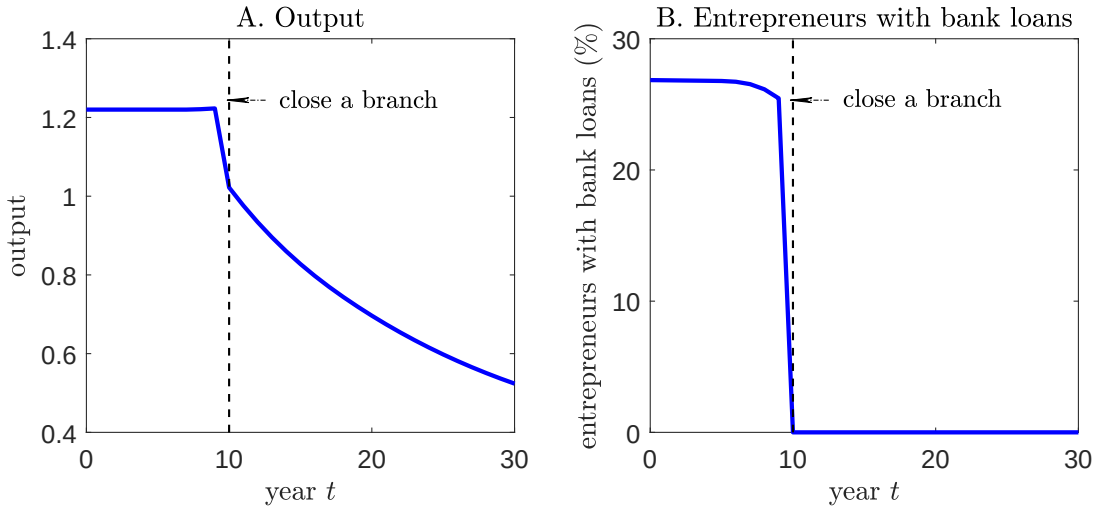


Figure OA.49: Number of branches opened and closed in each year from 1987 to 2011.

anticipating a future branch closure would have a negligible impact, if any, on households' current decisions and aggregate output. To visualize this, consider a market with median productivity ($Z_i = 0.99$) and a bank branch in our baseline specification, taking the parameter values and the path of nation-wide equilibrium interest rates as given by the baseline calibration. The market is in the steady state in year $t = 0$. Suppose that the branch is closed in year $t = 10$, which is fully anticipated since year $t = 0$. After the branch closure, we set $\psi_{i,t} = \xi_{i,t} = \infty$ for $t \geq 10$. We fully solve the transitional dynamics from $t = 0$ to $t = 30$ with the local equilibrium wage determined to clear the market's labor market. Figure OA.50 shows that the market's output (the sum of the outputs of all entrepreneurs in the market) and fraction of entrepreneurs with bank loans are virtually unchanged from year $t = 0$ to year $t = 9$.

Figure OA.51 shows that the transitional dynamics under the alternative specification are almost identical to the dynamics under our baseline specification for the 1986-1996 period. The GDP growth between 1986 and 1996 in the alternative specification is 26.9%, which is only slightly larger than the 26.5% GDP growth in the baseline specification, primarily because entrepreneurs anticipate future branch openings after 1996 and begin to accumulate wealth in advance. Unsurprisingly, the dynamics after 1996 diverge significantly between the specifications that do and do not allow new branches beyond our featured baseline period. For example, the GDP growth between 1986-2011 predicted by the baseline specification is 28.3% (blue solid line in panel A), whereas that predicted by the alternative specification is 32.3% (black dashed line in panel A) because of continued bank expansion after 1996.

The similar quantitative effects of the baseline specification and the alternative specification for 1986-1996 imply that the anticipation of future branch openings does not have a large quantitative

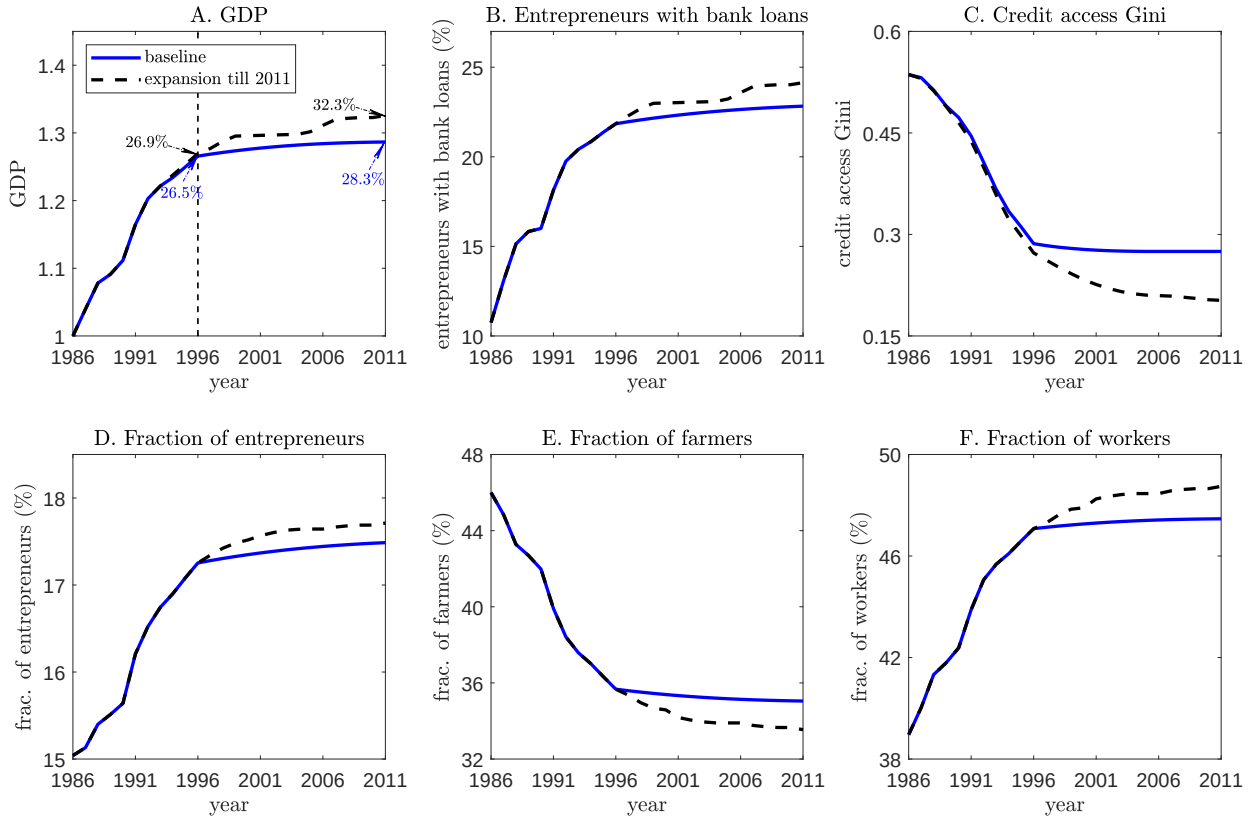


Note: We focus on the steady state of a single market with median productivity ($Z_i = 0.99$) and a bank branch at $t = 0$. The branch is closed at $t = 10$, which is fully anticipated since $t = 0$. After the branch closure, we set $\psi_{i,t} = \xi_{i,t} = \infty$ for $t \geq 10$. Panels A and B plot the market's output (the sum of the outputs of all entrepreneurs in the market) and fraction of entrepreneurs with bank loans from year $t = 0$ to year $t = 30$.

Figure OA.50: Illustration of the anticipation effect of a branch closure.

impact on current aggregate transitional dynamics. This is due to three reasons. First, focusing on a single market, the quantitative impact of opening a branch in the market in the future has a small impact on the market's current output. Figure OA.52 illustrates this point by plotting a single market's output from year $t = 0$ to year $t = 30$. We assume that the market has median productivity ($Z_i = 0.99$). The market is in a steady state at $t = 0$ and initially far away from existing branches (i.e., we set $\psi_{i,t} = \xi_{i,t} = \infty$ for $t < 10$). A new branch opens in the market at $t = 10$, which is fully anticipated since $t = 0$. It is clear that the output of the market remains flat until $t = 5$ and starts to increase only slightly from year $t = 6$, which is approximately 5 years from the branch-opening year ($t = 10$). Importantly, the market's output growth from year $t = 6$ to year $t = 9$ (most of the movement is in year $t = 9$) is much smaller than the growth after the new branch is opened in year $t = 10$.

Second, given that approximately 60% of the markets in the economy already have bank branches by 1996, the anticipation of future branch openings in some markets would not have a substantial impact, if any, on the aggregate dynamics, which reflects the weighted average output across all markets. Figure OA.52 shows that the anticipation effect only plays a role if the new branch is to be opened within the next 5 years. Thus, only the new branches to be opened between 1997 and 2001 will affect the aggregate output in 1996. In the data, the total number of new branches opened in this 5-year period is 119, much smaller than the total number of markets in the whole economy ($N = 1,428$).



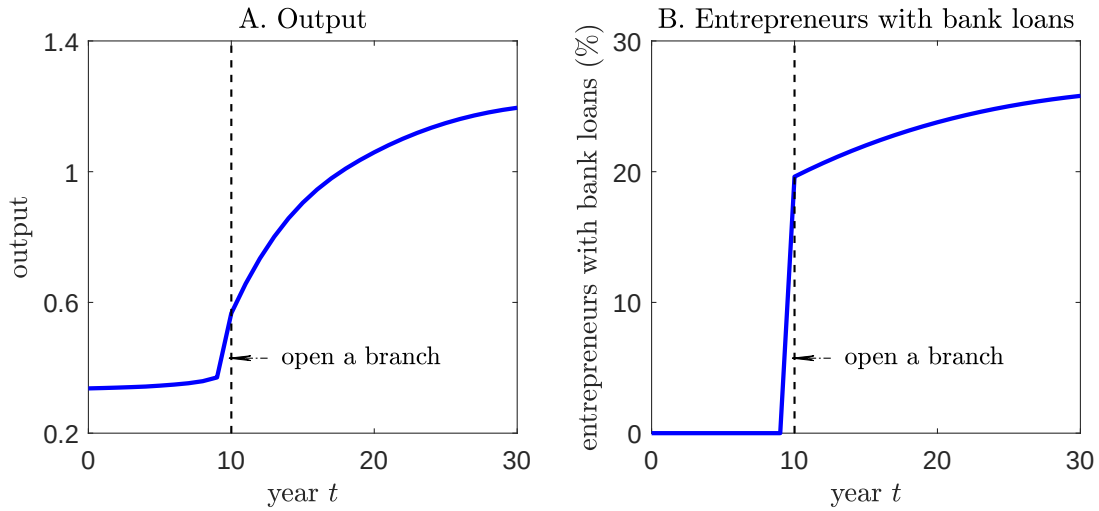
Note: The blue solid line represents the baseline specification in which new branches are opened only from 1987 to 1996, and the distribution of branches after 1996 is fixed. The black dashed line represents the alternative specification in which new branches continue to open from 1987 to 2011.

Figure OA.51: Comparison of the baseline and alternative specifications of bank expansion for the 1986-2011 period.

Third, the central authority will first open branches in more profitable markets in our model. These markets generally have larger market size $\Pi_{i,t}$ and productivity Z_i, meaning that they have larger weights in determining the economy's GDP. Thus, markets that do not have branches by 1996 carry much smaller weights than the markets that do. This further implies that the 119 new branches opened between 1997 and 2001 will have a relatively small impact on the economy's GDP because they are opened in markets with smaller weights.

4.8 The 1997 Financial Crisis

Our baseline model does not address the 1997 financial crisis. According to the World Bank Enterprise Survey, the GDP of Thailand declined significantly by 2.8% and 7.6% in 1997 and 1998, respectively. The Thai economy then recovered in 1999 with a positive GDP growth rate of 4.6%. In this appendix section, we show that the model-implied transitional dynamics between



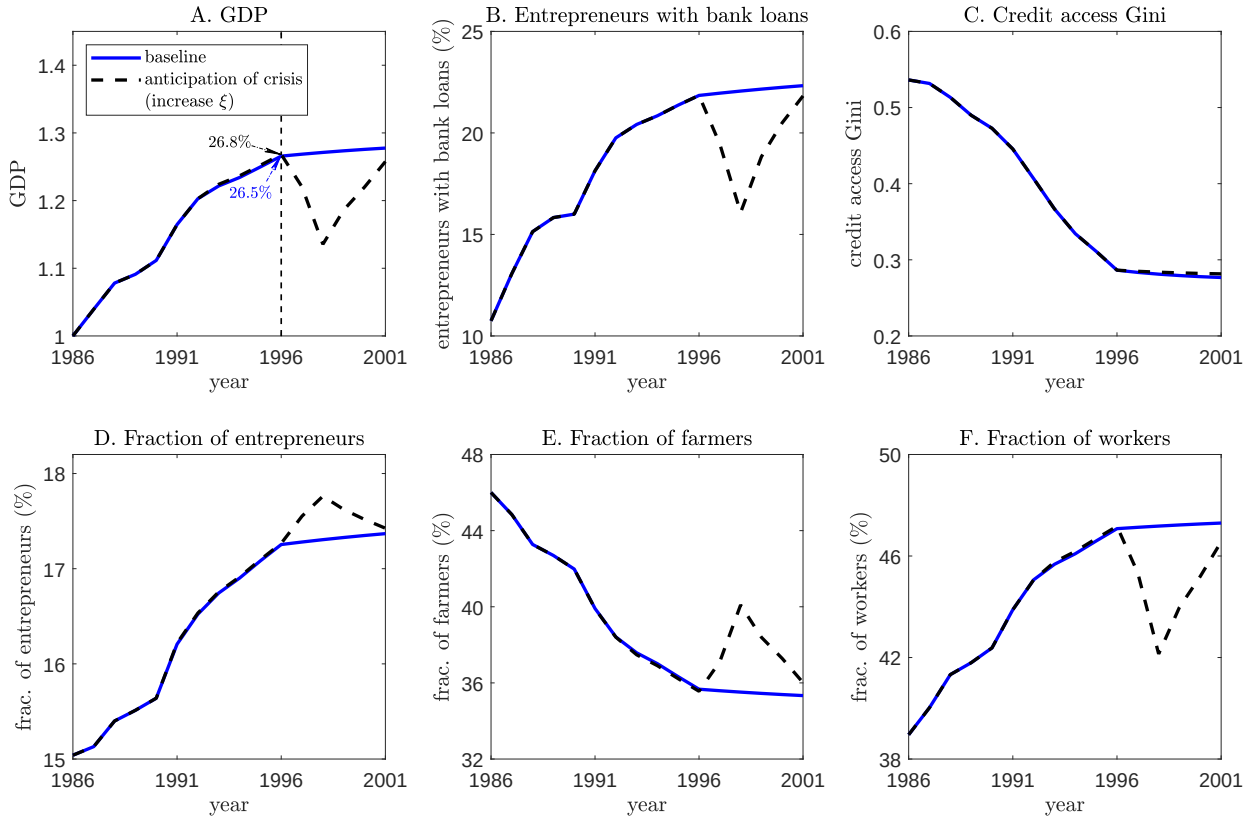
Note: We focus on a single market that is in a steady state at $t = 0$. The market has median productivity ($Z_i = 0.99$) and is far away from existing branches before $t = 10$ (i.e., we set $\psi_{i,t} = \xi_{i,t} = \infty$ for $t < 10$). There is a new branch opened in the market at $t = 10$. The new branch is fully anticipated since $t = 0$. The figure plots the market's output (the sum of the output of all entrepreneurs in the market) and fraction of entrepreneurs with bank loans from year $t = 0$ to year $t = 30$.

Figure OA.52: Illustration of the anticipation effect of opening a new branch.

1986 and 1996, the high growth and rapid bank expansion period of interest, is almost unchanged even if the financial crisis had been fully anticipated since the beginning of our simulation.

As our goal is not to offer a micro-founded explanation of the financial crisis, we model this event by simply introducing shocks to key plausible parameters of our model in 1997 and 1998. All of the other parameters are taken directly from our baseline calibration in Table 2 of the main text. To ensure that our results are robust, we consider three different experiments.

In the first experiment (Figure OA.53), we calibrate a higher liquidation loss rate ξ in 1997 and 1998 so that the model-implied GDP declines are 2.8% and 7.6% in these 2 years, respectively, consistent with the numbers in the data. A higher ξ captures a tightened borrowing constraint, which mimics a credit crunch. Figure OA.53 shows that compared to the baseline specification (blue solid line), anticipation of the financial crisis in 1997-1998 has a very small impact on transitional dynamics between 1986 and 1996. Our model implies that in the specification with the financial crisis (black dashed line), the GDP growth between 1986 and 1996 is 26.8%, which is higher than the 26.5% GDP growth in our baseline specification. This is because entrepreneurs anticipate the credit crunch and choose to accumulate more wealth in advance, which allows them to produce more before the financial crisis occurs. The model predicts that the financial crisis will generate a significant decline in the fraction of entrepreneurs with bank loans (panel B) as many entrepreneurs choose not to pay the credit entry costs when borrowing constraints are more stringent. The wage and interest rate will decrease during the crisis because of the decreases in capital and labor demand due to more stringent borrowing constraints. Panel D of Figure



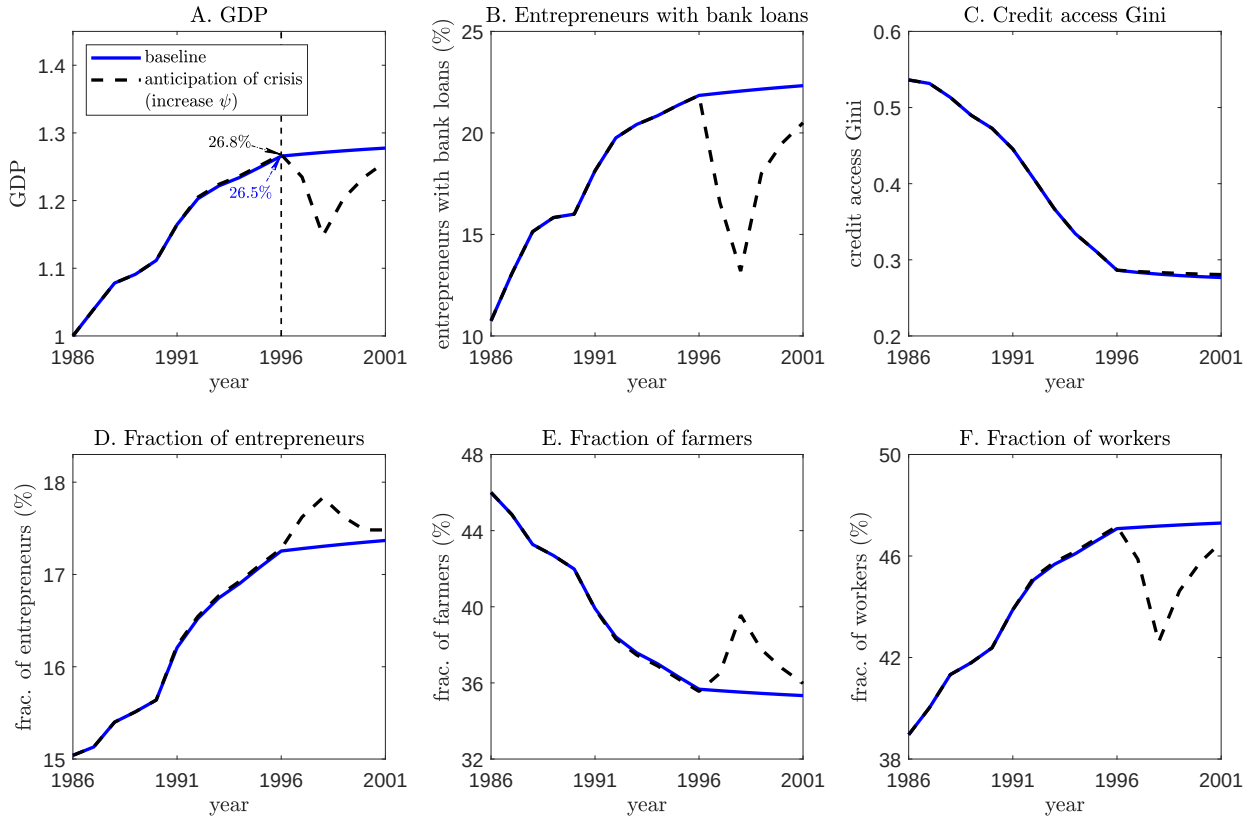
Note: We calibrate the value of ξ in 1997 and 1998 at 0.82 and 0.89, respectively, to generate 2.8% and 7.6% declines in GDP in 1997 and 1998. The value of ξ in other years is set at 0.59 according to our baseline calibration in Table 2 of the main text.

Figure OA.53: Using shocks to liquidation loss rate ξ to generate the financial crisis in 1997-1998.

OA.53 shows that the fraction of entrepreneurs increases during the crisis because the lower wage motivates some workers to become entrepreneurs. However, because each entrepreneur employs fewer workers on average due to the tightened borrowing constraint, the aggregate demand for workers falls, leading to an increase in the fraction of farmers (panel E) and a decrease in the fraction of workers (panel F). Panel C shows that the financial crisis, which is modelled to affect every market of the economy, does not change the credit access inequality across markets.

In the second experiment (Figure OA.54), we generate a credit crunch by increasing the calibrated value of ψ in 1997 and 1998. A higher ψ implies that all of the markets will experience uniformly higher credit entry costs. Figure OA.54 shows that this specification with the financial crisis also generates similar transitional dynamics as our baseline specification for the 1986-1996 period. The transitional dynamics after 1996 are similar to those in the first experiment (Figure OA.53), although with small differences in the magnitudes of the responses in various variables.

In the third experiment (Figure OA.55), we consider what would occur if the 1997-1998

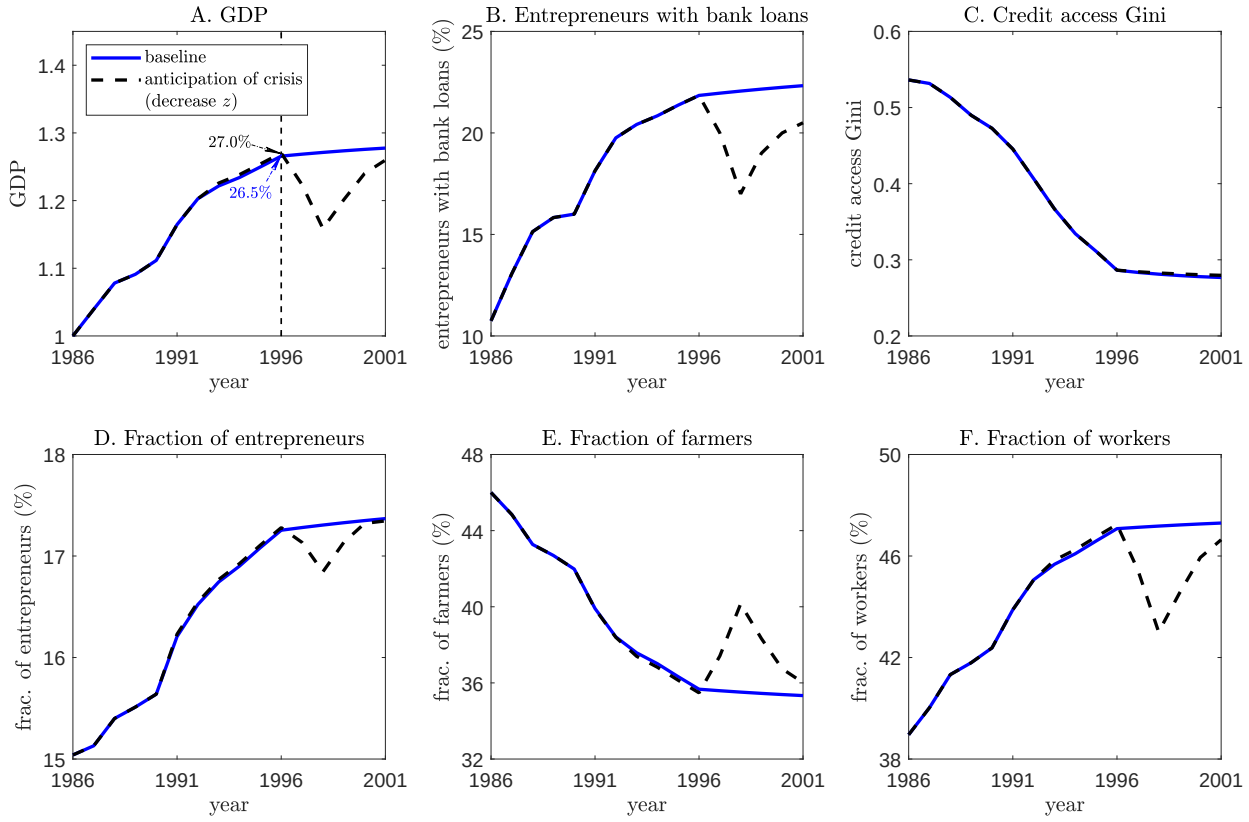


Note: We calibrate the value of ψ in 1997 and 1998 at 2 and 4, respectively, to generate 2.8% and 7.6% declines in GDP in 1997 and 1998. The value of ψ in other years is set at 0.83 according to our baseline calibration in Table 2 of the main text.

Figure OA.54: Using shocks to credit entry costs ψ to generate the financial crisis in 1997-1998.

crisis were caused by a productivity shock rather than a financial shock. We assume that all entrepreneurs' talent z_t is scaled by a factor of $\varepsilon_{z,t} < 1$ in 1997 and 1998. We calibrate $\varepsilon_{z,t}$ in each of the 2 years so that the model-implied negative GDP growth rates in 1997 and 1998 are identical to the numbers in the data. In all other years, we set $\varepsilon_{z,t} = 1$, as in our baseline specification without the scaling factor. Figure OA.55 shows that, again, this specification with the financial crisis generates similar transitional dynamics as our baseline specification for the 1986-1996 period.

However, the transitional dynamics of some variables after 1996 display different patterns than the first two experiments in Figures OA.53 and OA.54. Notably, the decline in the fraction of entrepreneurs with bank loans during the crisis in panel B of Figure OA.55 is much less than those in panel B of Figures OA.53 and OA.54. This is perhaps not surprising given that we are considering a negative shock to productivity in the third experiment, rather than a direct shock to the financial constraint. As explained above, the higher fraction of entrepreneurs in the first



Note: We uniformly scale the talent z of all entrepreneurs by a factor of $\varepsilon_{z,t} = 0.972$ and $\varepsilon_{z,t} = 0.93$ in 1997 and 1998, respectively, to generate 2.8% and 7.6% declines in GDP in 1997 and 1998.

Figure OA.55: Using shocks to household talent z to generate the financial crisis in 1997-1998.

two experiments is caused by the general equilibrium effect through lower wages. In the third experiment, the equilibrium wage is also lower during the crisis; however, its positive impact on the fraction of entrepreneurs is dominated by the negative impact of lower productivity. In other words, the negative productivity shock directly reduces the returns of businesses, exerting a significantly negative impact on entrepreneurial activities.

5 Numerical Algorithm

In this section of the appendix, we describe our numerical algorithm that solves the baseline model. The key difficulty lies in solving the central authority's location choice problem because of the interactions among markets in a network. This problem is NP-hard and essentially similar to the vertex covering problem (Karp, 1972). The profits of a branch can potentially be reduced by the establishment of other branches in nearby markets because households are allowed to travel

to other markets to make deposits. Solving this problem requires the central authority to know the profits arising from all of the possible combinations of new branch locations to attain the highest profits. The computational complexity of a brute-force algorithm is $O(N!)$.[26]

Because the problem is NP-hard, it is impossible to solve the model exactly. Thus, we propose a tractable algorithm that solves the model approximately. The fundamental design of our algorithm breaks the central authority's location choice problem across the country into a set of smaller problems, each of which focuses on specific segments of the country. These segments are constructed such that there is limited interaction among locations in different segments. This allows us to solve the problem of each segment separately, which is computationally tractable. Then, we solve the whole country's problem by aggregating all segments' solutions subject to certain feasibility constraints.[27]

After we solve the full model using the proposed approximation algorithm, we evaluate its performance at the province level in Section 5.6. We show that among the 44 provinces that have fewer than $100,000$ possible location choices, our approximation algorithm produces results identical to a brute-force algorithm that solves the model exactly. This increases our confidence in the algorithm.

The remainder of this section proceeds as follows. Subsection 5.1 provides an overview of our algorithm. Subsection 5.2 describes the algorithm that solves the initial distribution of households. Subsection 5.3 describes how the markets are partitioned into segments, and Subsection 5.4 describes the algorithm that solves the segment-level location choice. Subsection 5.5 describes the algorithm that aggregates segment-level location choice to obtain the optimal country-level location choice. Finally, Subsection 5.6 evaluates the performance of the algorithm at the province level.

[26]The main computational difficulty is caused by the spatial interaction among bank branches. If instead, we restrict households from traveling to other markets, the profits of different branches will not be interdependent, except through the general equilibrium effects. In this case, to open k branches, the central authority only needs to find the k markets that yield the highest profits individually without considering their interactions with other markets. This requires doing at most N calculations, so the computational complexity is $O(N)$.

[27]Zheng (2016) considers how retail chains choose store locations with spatially interdependent profits. Our algorithm is similar to that of Zheng (2016) in that we both apply a clustering algorithm to partition locations into multiple segments, such that the profits of locations in different segments are roughly independent. This allows us to solve smaller-scale problems. However, our algorithm is different because we further reduce the dimensionality of the problem within each segment based on economic intuition, and we transform the country-level problem into a tractable multidimensional multiple-choice knapsack problem. Our algorithm is related to those proposed by several other papers. For example, Jia (2008) considers a static combinatorial choice problem of how chain stores make entry decisions. By exploiting the supermodularity of the return function, Jia (2008) proposes an algorithm that reduces the set of potential combinations and computes the return of each remaining combination to obtain the optimal combination. Eckert and Arkolakis (2017) generalize the algorithm of Jia (2008) to problems with return functions satisfying single crossing differences. They reduce the set of potential combinations and check the remaining combinations recursively. Oberfield et al. (2020) consider how heterogeneous firms choose the number, size, and locations of their plants. They analytically solve a limit formulation of the problem where the number of plants per firm is large, such that firms choose the density of plants instead of a discrete number of plants.

5.1 Algorithm Overview

We provide an overview of the process of our algorithm below. We consider $t = 0$ as representing the first year of our study period, 1986, and thus $t = 1, ..., 10, ...$ represents 1987, ..., 1996, ... We proceed as follows.

(i) Solve the initial distribution of households, $\{\phi_0(s|i)\}_{i=1}^N$, which is the steady-state distribution of households corresponding to an economy with a branch distribution fixed at time $t = 0$ and no new branch openings for $t \geq 1$. See Online Appendix 5.2 for the algorithm.

(ii) Partition markets into G segments. See Online Appendix 5.3 for the algorithm.

(iii) Guess the interest rates $\{r_t\}_{t=0}^T$ and the wages $\{w_{i,t}\}_{t=0}^T$ for each market $i = 1, ..., N$, where we set $T = 200$ to ensure that the economy will eventually reach the steady state corresponding to the distribution of branches at time $t = 10$. Initialize $flag = 0$.

(iv) Given the guessed interest rates $\{r_t\}_{t=0}^T$ and wages $\{w_{i,t}\}_{t=0}^T$, for each segment $g = 1, ..., G$ defined in step (ii), we solve the optimal location-choice problem within the segment as a function of the number of branches opened in each year t in the segment. To compute the profits of opening a branch in market i, we need to solve households' optimal decisions corresponding to the choice of branch locations and conduct forward simulations to obtain the distribution of households $\{\phi_t(s|i)\}_{t=1}^T$. See Online Appendix 5.4 for the algorithm.

(v) Aggregate the potential location choices of all segments to formulate a multidimensional multiple-choice knapsack problem at the country level. Then, solve this problem to obtain the approximately optimal location choice for the whole country. See Online Appendix 5.5 for the algorithm.

(vi) Compute the implied equilibrium interest rates $\{\widehat{r}_t\}_{t=0}^T$ and wages $\{\widehat{w}_{i,t}\}_{t=0}^T$ for each market $i = 1, ..., N$ that clear the capital and labor markets (i.e., condition (iii) and (iv) in Section 3.4 of the main text).

(vii) Calculate $\text{diff}_r = \max\{|r_t - \widehat{r}_t| : t = 1, ...T\}$. If $\text{diff}_r > 10^{-5}$, set $flag = 1$ and update the initial guess for $\{r_t\}_{t=0}^T$ using $(\{r_t\}_{t=0}^T + \{\widehat{r}_t\}_{t=0}^T)/2$. For each market i, calculate $\text{diff}_{w,i} = \max\{|w_{i,t} - \widehat{w}_{i,t}| : t = 1, ...T\}$. If $\text{diff}_{w,i} > 10^{-5}$, $flag = 1$ and update the initial guess for $\{w_{i,t}\}_{t=0}^T$ using $(\{w_{i,t}\}_{t=0}^T + \{\widehat{w}_{i,t}\}_{t=0}^T)/2$.

(viii) If $flag = 1$ return to step (iv) with the updated initial guess and reset $flag = 0$. If $flag = 0$, the solution is found.

We provide further clarifications for step (iv) in the above algorithm. This step can be implemented in a straightforward manner when the model does not allow for migration. In this

case, the optimal choice of branch locations and household decisions in each segment g do not depend on the locations of branches in other segments, given the guessed interest rates and wages. However, when there is migration, the decisions of households in segment g depend on the locations of branches in other segments. Thus, to solve the problem for segment g, we first need to guess the optimal branch locations in other segments.

An informative guess can be obtained using the optimal branch locations corresponding to an alternative model specification without migration, which is exactly the calibration with $\kappa = \infty$ and $\eta = 0$ analyzed in Online Appendix 4.1.2. Thus, to solve the baseline model with migration, we need to first solve an alternative model specification without migration using the above algorithm. Then, we can take the solution for the distribution of branch locations as an informative guess for the branch locations in the baseline model and solve the baseline model using the above algorithm.

5.2 Initial Distribution

We solve the steady state corresponding to the distribution of branch locations in 1986, which is used as the initial distribution for step (i) of the algorithm in Online Appendix 5.1. We proceed as follows.

(i) Specify an arbitrary initial distribution of households, $\{\phi_0(s|i)\}_{i=1}^N$.

(ii) Guess the interest rate r_{ss} in the steady state.

(iii) Guess the wages $\{w_{i,ss}\}_{i=1}^N$ in the steady state.

(iv) For each market $i = 1, ..., N$:

  (a) Given the interest rate r_{ss} and wage $w_{i,ss}$, solve the household problem (equations (3.5)-(3.11) of the main text and (3.1)-(3.4) of Online Appendix 3.1) in the absence of migration to obtain the policy functions for consumption, cash, deposits, portfolio adjustments, occupation choices, credit access, and the demands for capital and labor.

  (b) Starting from time $t = 0$, simulate market i for T years using the policy functions solved in step (a) and the initial distribution $\phi_0(s|i)$ in step (i) to obtain the steady-state distribution of households $\phi_T(s|i)$. We set $T = 200$, which is a sufficiently long period for the economy to reach the steady state.

  (c) Use the steady-state distribution $\phi_T(s|i)$ to solve the implied wage $\widehat{w}_{i,ss}$ that is consistent with the labor market-clearing condition (i.e., condition (iv)) in Section 3.4 of the main text. We first calculate the excess labor supply when the wage is f. If the excess labor supply is positive, we set $\widehat{w}_{i,ss} = f$, indicating that farmers exist in the

economy. If the excess labor supply is negative, we find a value of $\widehat{w}_{i,ss} > f$ that clears the labor market according to equation (3.20) of the main text. Finally, we calculate $\text{diff}_{w,i} = |w_{i,ss} - \widehat{w}_{i,ss}|$. If $\text{diff}_{w,i} > 10^{-5}$, return to step (iv, a) using $(w_{i,ss} + \widehat{w}_{i,ss})/2$ as the initial guess for the wage.

(v) Use the steady-state distribution $\{\phi_T(s|i)\}_{i=1}^N$ and wages $\{w_{i,ss}\}_{i=1}^N$ solved in step (iv) and the initial market size $\{\Pi_{i,0}\}_{i=1}^N$ to solve the implied interest rate $\widehat{r}_{ss}$ that clears the capital market according to equation (3.18) of the main text. Calculate $\text{diff}_r = |r_{ss} - \widehat{r}_{ss}|$. If $\text{diff}_r > 10^{-5}$, return to step (iv) using $(r_{ss} + \widehat{r}_{ss})/2$ as the initial guess for the interest rate and using $\{w_{i,ss}\}_{i=1}^N$ (solved in step (iv)) as the initial guess for wages.

(vi) The steady-state distribution $\{\phi_T(s|i)\}_{i=1}^N$ is used as the initial distribution of households for step (i) of the algorithm in Online Appendix 5.1.

Now, we discuss the grids used to discretize households' state space. Households' occupation choices and portfolio adjustments make their value functions nonconvex in certain regions. To mitigate this problem and improve the accuracy of our numerical solutions, we solve value functions in terms of $s' = \{z, m + a, m/(m + a)\} = \{z, b, \iota\}$ rather than $s = \{z, m, a\}$, as defined in the main text.

We use collocation methods to solve households' value and policy functions. Let $S_z \times S_b \times S_\iota$ be the grid of collocation nodes for households' value and policy functions, where $S_z = \{z_1, ..., z_{n_z}\}$, $S_b = \{b_1, ..., b_{n_b}\}$, and $S_\iota = \{\iota_1, ..., \iota_{n_\iota}\}$. Grids in each dimension of the state vector are taken with equal spacing. Talent z is discretized into $n_z = 11$ grids from 1 to 2.7, which is the value corresponding to the 99th percentile of the cumulative distribution function. Total wealth b is discretized into $n_b = 201$ grids from 0 to 200. The cash-wealth ratio ι is discretized into $n_\iota = 11$ grids from 0 to 1. When conducting simulations, we track the probability density functions of households rather than a large number of households. We use finer grids for simulations. Total wealth is discretized into $10,001$ grids from 0 to 200, and the cash-wealth ratio is discretized into 21 grids from 0 to 1. The discretization of talent is unchanged.

5.3 Algorithm for Partitioning Markets into Segments

The goal of our partitioning algorithm is to ensure that markets in different segments are far away from each other, although markets within the same segment are allowed to be close to each other. This property implies that few households will travel across segments to obtain financial services. Thus, we can assume that the profits of a branch only depend on the distribution of branches within the same segment and not on branches in other segments.

We use the k-medoids clustering algorithm in machine learning (Kaufman and Rousseeuw,

2005) to partition $N = 1,428$ markets into $G = 35$ segments, indexed by $g = 1, ..., G$. The algorithm is as follows:

(i) Arbitrarily choose G markets as medoids and label them $m_1, ..., m_G$. Denote by $\mathbb{M}$ the set of these medoids.

(ii) Based on the distances τ_{ij} for $i, j = 1, ..., N$, find the closest medoid for each market. Define segment g as the set of markets for which the closest medoid is $m_g \in \mathbb{M}$.

(iii) In each segment, find a new medoid that corresponds to the market that minimizes the sum of its distance from every other market within the same segment. Denote by $\mathbb{M}'$ the updated set of medoids.

(iv) Stop if $\mathbb{M}' = \mathbb{M}$. Otherwise, return to step (ii) with the set of medoids $\mathbb{M}'$.

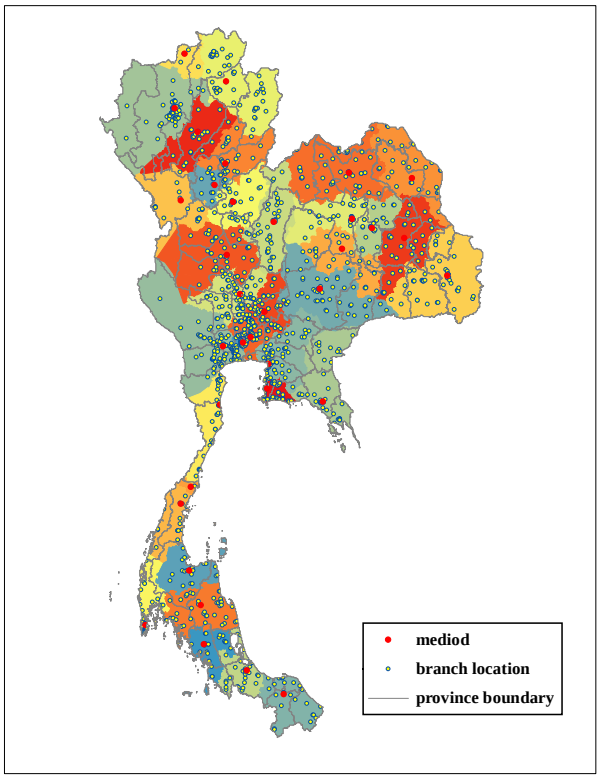
The partition outcome largely depends on the initial choice of medoids. Thus, we consider 500,000 independent random initial choices in step (i) and choose the partition that minimizes the following loss function:

$$L = \sum_{g=1}^{G} \mathbb{1}_{\{\min\{\tau_{ij}: i \in g, j \notin g\} < \overline{D}\}}.$$

The loss function counts the number of segments whose distance from the nearest segment is shorter than $\overline{D}$. A larger $\overline{D}$ imposes a greater penalty on the partition that yields segments closer to each other. The value of $\overline{D}$ must be chosen consistently with the number of segments G because the segments will naturally become less separate as the number of segments increases. We set $G = 35$ and $\overline{D} = 30$ so that each segment has a small number of markets and the segments are sufficiently distant from each other.

Figure OA.56 illustrates the segments of our partition. In general, we find that the boundaries of the segments are close to the boundaries of the provinces. Panel A of Figure OA.57 shows that within each segment, the number of potential locations for new branches (i.e., markets without a branch in 1986) ranges from 2 to 75. Panel B of Figure OA.57 shows that across segments, the distance from the nearest segment ranges from 27.3 minutes to 84.9 minutes (based on car travel time), and that most segments are more than 35 minutes' drive away from their neighbors. In our calibration, households choose not to borrow when the distance from the nearest bank branch is greater than 27 minutes (panel A of Figure OA.7). Moreover, households choose not to make deposits when the distance from the nearest bank branch is greater than 45 minutes (panel B of Figure OA.7), and the deposit-cash ratio is as low as 5% when the distance from the nearest bank branch is approximately 27 minutes. Given that the smallest distance between segments is 27.3 minutes, it seems to be a reasonable approximation to ignore the interdependence of profits between markets in different segments.[28]

[28] After we solve the full model, we double-check and confirm that it is not optimal to force the central authority



Note: This figure plots the segments of our partition with $G = 35$ and $\overline{D} = 30$. Segments are indicated by different colors. The red dots represent the 35 medoids, the yellow dots represent branch locations, and the gray lines represent the province boundaries.

Figure OA.56: Partitioning of markets into segments.

5.4 Segment-Level Location Choice

This section describes the algorithm that solves the segment-level location choice. Because the largest segment contains 75 potential locations for new branches (see panel A of Figure OA.57), the segment-level location choice problem remains fairly untractable. Thus, we first reduce the dimensionality of each segment based on economic intuition. Then, we find all potential location choices in each segment that could constitute the optimal location choice at the country level.

5.4.1 Reduce the Dimensionality Within Each Segment

Within each segment g, we classify the potential locations (i.e., markets) into two disjoint groups with the following property. The profits of a branch opened in a market in group 1 are not

to open a pair of branches that are close to each other (i.e., less than 45 minutes apart) but lie in two neighborhood segments.

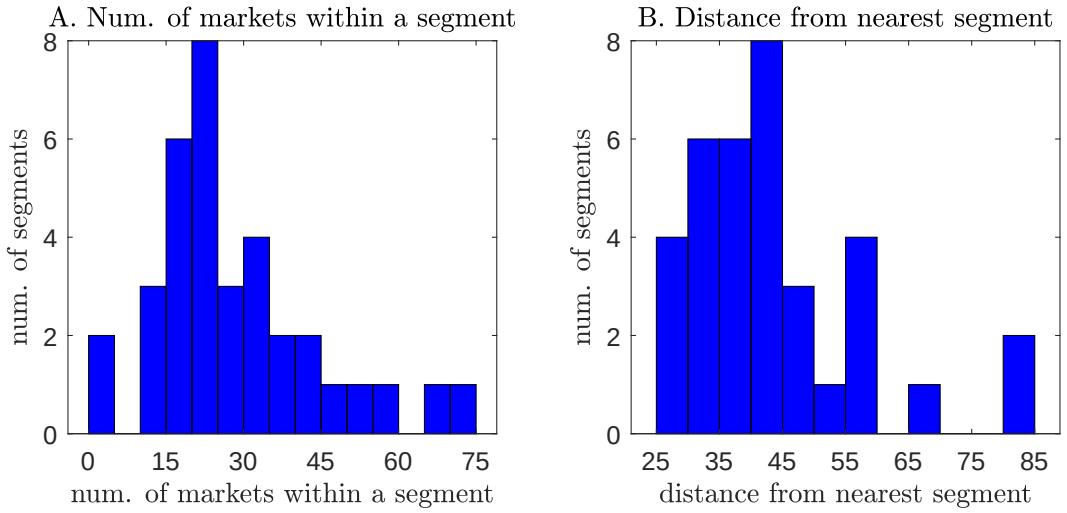


Figure OA.57: The distribution of the number of markets within a segment and distance from the nearest segment.

affected by the existence of any other branch in the segment. By contrast, the profits of a branch opened in a potential location in group 2 may depend on whether there are branches in other markets in the segment. As we show below, this dichotomy greatly reduces the dimensionality of our location choice within the segment because all of the markets in group 1 can be treated individually without considering their potential interactions with other markets when determining the profits from opening a branch.

We assign any market in the segment to group 1 if the market's distance from any other market in the segment is greater than 50 minutes or if the market's population density is higher than the branch's capacity h. The profits from opening a branch in such markets will not be affected by the existence of branches in other markets because markets belonging to the former case are far away from other markets. For a market in the latter case, households from other markets are restricted from traveling to this market due to the market's high population density, according to equation (3.1) of the main text. The markets not assigned to group 1 are all assigned to group 2.

Panel A of Figure OA.58 shows that within each segment, the number of potential locations in group 1 ranges from 0 to 44. Panel B shows that within each segment, the number of potential locations in group 2 ranges from 0 to 34. This greatly reduces the dimensionality of the segment-level location choice because now we only need to consider possible combinations of 34 markets in the largest segment, rather than 75 markets (as indicated by panel A of Figure OA.57).

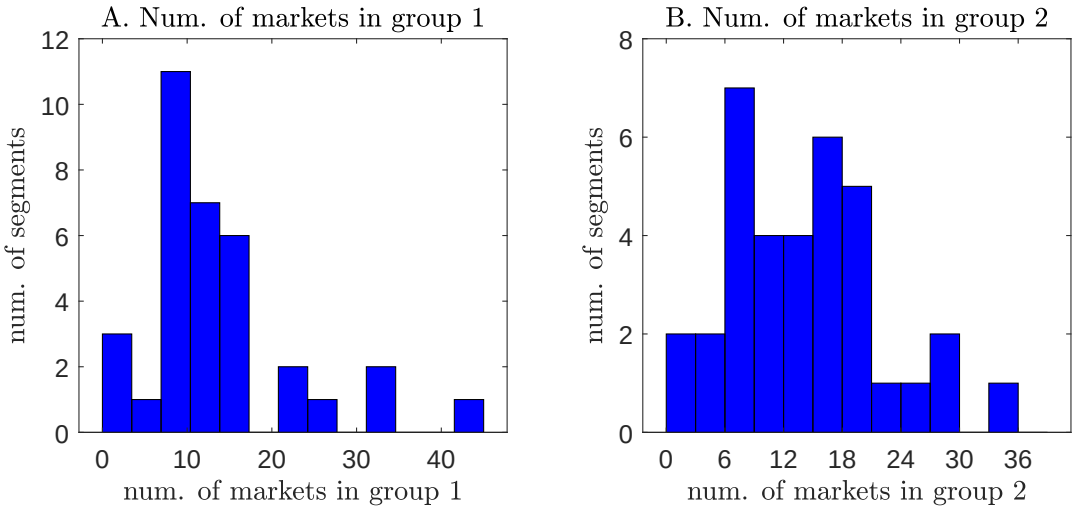


Figure OA.58: The distribution of the numbers of markets in groups 1 and 2.

5.4.2 Location Choice Within Each Segment

Based on the classification above, we solve the location choice within each segment g. Note that our eventual goal is to find the optimal location choice for the whole country, which requires us to solve two things for each segment $g = 1, ..., G$. First, we need to solve the number of branches $l_{g,t}$ opened in each year $t = 1, ..., T$ for segment g, subject to the constraint $\sum_{g=1}^{G} l_{g,t} = n_t$, where $t = 1$ represents year 1987, $t = T$ represents year 1996, and n_t is total number of branches opened in year t, exogenously given by the data. Second, we need to solve where to open the $l_{g,t}$ branches in year t within segment g.

Denote by $\left\{l^*_{g,t}\right\}_{t=1}^{T}$ the optimal number of branches opened in each year t for segment g. Because we do not know $\left\{l^*_{g,t}\right\}_{t=1}^{T}$ ex-ante, for each segment g, we need to solve the optimal profits and location choices for a range of $l_{g,t}$ in each year t. These location choices constitute the solution space for the optimization problem at the country level (see Online Appendix 5.5).

Of course, it is infeasible to consider all possible values of $\{l_{g,t}\}_{t=1}^{T}$ because the space is too large. Thus, we have to restrict the range of $l_{g,t}$ in each year t. We find that it is sufficient to consider the following range.[29] If there are $\{n_{g,t}\}_{t=1}^{T}$ new branches opened in segment g in the data, we consider all possible $\{l_{g,t}\}_{t=1}^{T}$ that satisfy two conditions. First, we restrict the total number of branches opened in segment g from $t = 1$ to $t = T$ by $\max\{\sum_{t=1}^{T} n_{g,t} - 10, 0\} \leq \sum_{t=1}^{T} l_{g,t} \leq \sum_{t=1}^{T} n_{g,t} + 10$. Second, we restrict the number of branches opened in segment g and year t by $\max\{n_{g,t} - 4, 0\} \leq l_{g,t} \leq n_{g,t} + 4$. Denote by L_g the set of such possible choices of $\{l_{g,t}\}_{t=1}^{T}$.

Next, we describe how to solve the optimal profits and location choices for all possible choices

[29] We conduct a robustness check by changing the condition to $\max\{\sum_{t=1}^{T} n_{g,t} - 15, 0\} \leq \sum_{t=1}^{T} l_{g,t} \leq \sum_{t=1}^{T} n_{g,t} + 15$ and $\max\{n_{g,t} - 5, 0\} \leq l_{g,t} \leq n_{g,t} + 5$, and find that the location choices after aggregating to the country-level are the same.

of $\{l_{g,t}\}_{t=1}^{T} \in L_g$. We propose a fast method by taking advantage of the two-group classification described in Online Appendix 5.4.1.

Highest-Possible 1-Year Net Profit. We first calculate the "highest-possible 1-year net profit" that a branch opened in year $t = 1$ can generate in each potential location. Considering potential location i as an example, the highest-possible 1-year net profit of opening a branch in location i is computed in three steps: 1) we compute the current-year profits generated by the branch opened in location i when no other new branches are opened in the segment. This gives the highest-possible gross profits of opening a branch in location i because opening new branches in other locations can potentially reduce the profits of location i due to cannibalization; 2) we compute the reduction in the current-year profits generated by branches that already exist at time $t = 0$ in segment g.[30] Note that the reduction is 0 if location i is far away from existing branches due to the absence of cannibalization; and 3) we compute the highest-possible 1-year net profit as the net increase in total profits of segment g after opening a branch in location i, which is equal to the highest-possible gross profits of opening a branch in location i (computed in step 1) minus the reduced profits of existing branches (computed in step 2).

Note that when computing the profits of a branch above, we also need to solve the decisions of households over time because households are forward looking. We use the algorithm of Buera and Shin (2013) to solve households' transitional dynamics, based on the discretized grids described in Online Appendix 5.2.

Solving the highest-possible 1-year net profit is time consuming because households are forward-looking. However, we note that the highest-possible 1-year net profit only needs to be solved once for each market, given a particular guess for the interest rates and wages (i.e., step (iii) of the algorithm in Online Appendix 5.1). That is, the same highest-possible 1-year net profit for markets in segment g applies to any choice of $\{l_{g,t}\}_{t=1}^{T} \in L_g$ in segment g. The reason is that the highest-possible 1-year net profit is defined to capture the net profits generated by opening one single branch only, which by definition, does not depend on the number of branches $l_{g,t}$ opened in each year t in segment g. This property makes it highly tractable to solve the optimal profits and location choices across all possible choices of $\{l_{g,t}\}_{t=1}^{T} \in L_g$, despite the potentially large number of elements in the set L_g. As we describe below, once we have the highest-possible 1-year net profit, the algorithm that solves the optimal location choice within the segment for a particular choice of $\{l_{g,t}\}_{t=1}^{T}$ boils down to basic arithmetic, which can be solved in microseconds.

Optimal Location Choice within the Segment. We sort all locations in the segment based on their highest-possible 1-year net profits from highest to lowest. There are two cases. First, if

[30] The purpose of step 2) is to account for the second term in equation (3.14) of the main text.

the top $\sum_{t=1}^{T} l_{g,t}$ locations are all in group 1, then these $\sum_{t=1}^{T} l_{g,t}$ locations are optimal in terms of maximizing joint profits because the profits from opening branches in the markets of group 1 do not depend on whether there are new branches opened in other locations within the segment. Moreover, due to the absence of cannibalization from opening new branches in other locations, the rank of 1-year profits is persistent. Thus, it is optimal to first open branches in locations at the top of the list and go down the list one by one. This is obviously numerically tractable.

Second, if some of the top $\sum_{t=1}^{T} l_{g,t}$ locations belong to group 2, then we need to search for additional combinations because the profits of the markets in group 2 could be affected by whether there are branches opened in nearby markets.[31] To give a specific example, suppose that among the top $\sum_{t=1}^{T} l_{g,t}$ locations with the highest-possible 1-year net profits, x locations are in group 1. Then we need to find $\sum_{t=1}^{T} l_{g,t} - x$ more markets by comprehensively searching for all combinations of size $\sum_{t=1}^{T} l_{g,t} - x$ among the other potential locations (including the rest of the locations in group 1 and all of the locations in group 2) to find the one that yields the highest increase in the total profits of segment g, which is equal to the joint profits generated by opening $\sum_{t=1}^{T} l_{g,t} - x$ branches net of the reduction in profits for branches that already exist at time $t = 0$ in segment g.[32]

The second case is essentially a submodular maximization problem subject to a knapsack constraint. It is numerically tractable for three reasons. First, conditional on the total number of new branches, $\sum_{t=1}^{T} l_{g,t}$, opened in segment g, we only need to solve the combinatorial programming problem once regardless of the split of $l_{g,t}$ across different year t. Second, under our calibration, x is generally close to $\sum_{t=1}^{T} l_{g,t}$.[33] Thus, exhaustively searching for all combinations is numerically tractable. Third, when solving for these combinations, we need to solve location-specific household problems, which could be time consuming because there are many combinations of locations. To speed up the code, we first solve household problems based on discretized grids of the distance from the nearest branch in each year.[34] Then, we use linear interpolation to obtain location-specific household decisions.

[31] In other words, the highest-possible profits associated with the locations in group 2 may not be attainable when other markets also have branches. By contrast, the highest-possible profits associated with the locations in group 1 are always attainable, regardless of whether other markets have branches.

[32] The solution may not be optimal because a branch opened in one of the x locations may reduce the profits of the remaining $\sum_{t=1}^{T} l_{g,t} - x$ locations due to cannibalization. We conduct a robustness check by only choosing the top $x - 1$ locations from group 1 and comprehensively searching for all combinations of size $\sum_{t=1}^{T} l_{g,t} - (x - 1)$ among other potential locations. The results are unchanged.

[33] This occurs for two reasons. First, group 1 contains all markets for which size exceeds capacity h. Opening a branch in these markets generally yields greater profits than opening a branch in other markets because of their large size. Second, although a branch opened in the markets belonging to group 2 can serve the home and nearby markets, under our calibration, the profits generated from the former are usually more important than the latter because the demand for loans decreases significantly when $d_{i,t}$ increases from 0 to 10 minutes (see panel C of Figure OA.12).

[34] Given a particular guess for the interest rates and wages (i.e., in step (iii) of the algorithm in Online Appendix 5.1), the household problem here only needs to be solved once and can be used for every segment g.

5.5 Aggregation and Country-Level Location Choice

For each segment $g = 1, ..., G$, our algorithm in Online Appendix 5.4 solves the optimal locations of opening $l_{g,t}$ new branches in year t to maximize the joint profits of all new branches opened from time $t = 1$ to time $t = T$ for all choices of $\{l_{g,t}\}_{t=1}^{T} \in L_g$.

Denote by $\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$ the set of branch locations of the $l_{g,t}$ branches opened in year t and segment g, corresponding to the choice of $\{l_{g,t}\}_{t=1}^{T} \in L_g$.[35] Denote by $\Theta_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$ the joint profits of the $l_{g,t}$ branches opened in year t and segment g, corresponding to the location choice $\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$. Thus, $\sum_{t=1}^{T}\Theta_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$ represents the total profits of all $\{l_{g,t}\}_{t=1}^{T}$ branches corresponding to the location choice $\left\{\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})\right\}_{t=1}^{T}$ in segment g,.

Now, we aggregate the above segment-level location choices to solve the central authority's optimal location choice at the country level. The central authority maximizes the total profits from opening $\{n_t\}_{t=1}^{T}$ branches subject to the constraint that the sum of the number of branches opened in each segment is equal to n_t in year t, as follows:[36]

$$\max_{\{\{l_{g,t}\}_{t=1}^{T}\in L_g\}_{g=1}^{G}} \sum_{g=1}^{G}\left[\sum_{t=1}^{T}\Theta_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})\right], \tag{5.1}$$

$$\text{s.t.} \quad \sum_{g=1}^{G} l_{g,t} = n_t \quad \text{for} \quad t = 1, ..., T. \tag{5.2}$$

Problem (5.1) is essentially a multidimensional multiple-choice knapsack problem. Items of different sizes ($\{l_{g,t}\}_{t=1}^{T}$) and values ($\sum_{t=1}^{T}\Theta_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$) are partitioned into disjoint classes (g). The aim is to pack exactly one item from each class to maximize the total value of packed items ($\sum_{g=1}^{G}[\sum_{t=1}^{T}\Theta_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})]$), subject to multiple knapsack constraints (5.2).

This problem can be solved exactly by the recursive algorithm proposed by Dudzinski and Walukiewicz (1987) in psuedopolynomial time. Denote by $\Phi^k(\{x_t\}_{t=1}^{T})$ the highest total profits from opening x_t branches in each year t in the first k segments with $x_t = 0, ..., n_t$ and $k = 1, ..., G$.

[35] Obviously, the set $\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$ is a function of the number of branches $l_{g,t}$ in year t and segment g because $l_{g,t}$ directly determines the cardinality of the set, i.e., $|\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})| = l_{g,t}$. The set $\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$ also depends on the number of branches, $\{l_{g,\tau}\}_{\tau\neq t}$, opened in other years in segment g, because to obtain $\Lambda_{g,t}(l_{g,t}|\{l_{g,\tau}\}_{\tau\neq t})$, our algorithm in Online Appendix 5.4 optimizes the joint profits of opening $\{l_{g,t}\}_{t=1}^{T} \in L_g$ new branches. In other words, conditional on the same $l_{g,t}$, if more or fewer branches are opened in other years $\tau \neq t$, the optimal solution for the locations of these $l_{g,t}$ opened in year t may change both because of cannibalization across branches and because the number of potential locations where new branches could open is given.

[36] Equations (5.1) and (5.2) are transformed from equations (3.14) and (3.15) of the main text.

Similar to equations (5.1) and (5.2), $\Phi^k(\{x_t\}_{t=1}^T)$ is defined as follows:

$$\Phi^k(\{x_t\}_{t=1}^T) = \max_{\{\{l_{g,t}\}_{t=1}^T \in L_g\}_{g=1}^k} \sum_{g=1}^k \left[\sum_{t=1}^T \Theta_{g,t}(l_{g,t}|\,\{l_{g,\tau}\}_{\tau\neq t})\right], \tag{5.3}$$

$$\text{s.t.} \quad \sum_{g=1}^k l_{g,t} = x_t \quad \text{for} \quad t = 1, ..., T. \tag{5.4}$$

We set $\Phi^k(\{x_t\}_{t=1}^T) = -\infty$ if the feasible region defined by equation (5.4) is empty. We set $\Phi^0(\{x_t\}_{t=1}^T) = 0$ for $x_t = 0, ..., n_t$ and compute Φ^k for $k = 1, ..., G$ in sequence by the following recursive equation:

$$\Phi^{k+1}(\{x_t\}_{t=1}^T) = \max_{\{l_{k+1,t}\}_{t=1}^T \in L_{k+1}} \Phi^k(\{x_t - l_{k+1,t}\}_{t=1}^T) + \sum_{t=1}^T \Theta_{k+1,t}(l_{k+1,t}|\,\{l_{k+1,\tau}\}_{\tau\neq t}). \tag{5.5}$$

Equation (5.5) maximizes the total profits from opening $\{x_t\}_{t=1}^T$ branches in the first $k+1$ segments. The optimal profits from opening n_t branches in each year t are given by $\Phi^G(\{n_t\}_{t=1}^T)$ and the optimal location choice can be easily backed out using the above recursive procedures.

The above algorithm needs to compute $\Phi^k(\{x_t\}_{t=1}^T)$ for all $\{x_t\}_{t=1}^T$, with $x_t = 0, ..., n_t$ and $k = 0, ..., G$. One standard way to exactly solve this is to use a branch and bound algorithm. However, the number of feasible values for the state vector $\{x_t\}_{t=1}^T$ is $\prod_{t=1}^T n_t \approx 6.9 \times 10^{15}$. Due to the high dimensionality of the state vector $\{x_t\}_{t=1}^T$, the branch and bound algorithm is not computationally tractable. Thus, the standard way to solve this high-dimensional multiple-choice knapsack problem is to use a heuristic algorithm (Parra-Hernandez and Dimopoulos, 2005; Yu, Zhang and Lin, 2007; Ren, Feng and Zhang, 2012). Our heuristic algorithm is essentially a greedy algorithm that solves the central authority's location choice problem year by year. The accuracy of this algorithm depends on the extent to which branches opened in different years interact with each other in determining the total profits of all new branches. After presenting the algorithm, we discuss and verify that this greedy algorithm yields a very good approximation for optimal branch locations in our setting. The algorithm is as follows:

First, we solve the locations of new branches opened at time $t = 1$ to maximize the profits generated by these branches. This simply involves solving the following one-dimensional multiple-choice knapsack problem:

$$\max_{\{\{l_{g,t}\}_{t=1}^T \in L_g\}_{g=1}^G} \sum_{g=1}^G \Theta_{g,1}(l_{g,1}|\,\{l_{g,\tau}\}_{\tau\neq 1}), \tag{5.6}$$

$$\text{s.t.} \quad \sum_{g=1}^G l_{g,1} = n_1. \tag{5.7}$$

Problem (5.6) can be solved by the recursive algorithm above (see equations (5.3)-(5.5)) except for modifying the objective function and knapsack constraint to focus on time $t = 1$ and replacing the T-dimensional state vector $\{x_t\}_{t=1}^T$ with a single state variable $x_1 = 0, ..., n_1$, which is computationally tractable.

After solving problem (5.6), we obtain the optimal number of branches $l_{g,1}^*$ for $g = 1, ..., G$ and the corresponding optimal locations $\Lambda_{g,1}^*$. We define that $L_g^{(1)} = \left\{\{l_{g,t}\}_{t=1}^T : \{l_{g,t}\}_{t=1}^T \in L_g, l_{g,1} = l_{g,1}^*\right\}$ for segment g. Note that $L_g^{(1)}$ defines the feasible set for the number of branches opened at time $t \geq 2$, taking both the number $l_{g,1}^*$ and locations $\Lambda_{g,1}^*$ of branches opened at time $t = 1$ as given. Our definition for $L_g^{(1)}$ does not explicitly state the constraint on locations $\Lambda_{g,1}^*$, but one should keep in mind that there is one such constraint.[37] That is, the locations of new branches opened at time $t \geq 2$ cannot overlap with $\Lambda_{g,1}^*$. Then, we solve the locations of new branches opened at time $t = 2$ as follows:

$$\max_{\{\{l_{g,t}\}_{t=1}^T \in L_g^{(1)}\}_{g=1}^G} \sum_{g=1}^G \Theta_{g,2}(l_{g,2} | \{l_{g,\tau}\}_{\tau \neq 2}), \tag{5.8}$$

$$\text{s.t.} \quad \sum_{g=1}^G l_{g,2} = n_2. \tag{5.9}$$

Thus, we obtain the optimal number of branches $l_{g,2}^*$ for $g = 1, ..., G$ and the corresponding optimal locations $\Lambda_{g,2}^*$. We continue the above procedure and solve the locations of new branches at time $t = 3, .., T$.

Because our heuristic algorithm solves the central authority's problem year by year using the exact algorithm of Dudzinski and Walukiewicz (1987), there is no approximation error for branches opened in the same year. Thus, the approximation error only depends on the extent to which jointly choosing branch locations across different years increases the total profits of new branches. There are few cases that would make the choice of branch locations different when they are jointly chosen across years. For one such illustrative case, consider three major markets, labeled M_1, M_2, and M_3, and multiple minor markets. Suppose that M_1 and M_2 are peripheral markets far away from each other, but both are close to the central market M_3. Suppose that a branch opened in M_1 and M_2 can obtain a one-time profit of π by itself (i.e., without including the profits of households from other markets) and a branch opened in M_3 can obtain a profit of $\pi' < \pi$ by itself. A branch opened in a minor market obtains negligible profit. Furthermore, suppose that a branch can obtain half of the profits from a nearby market that does not have a branch.

Now, consider a central authority that needs to open exactly one branch at time $t = 1$ and

[37] When implementing the algorithm, it is generally the case that the locations $\Lambda_{g,1}^*$ are uniquely determined given $l_{g,1}^*$, even though in principle, $\Lambda_{g,1}^*$ may also depend on the number of branches opened at time $t \geq 2$.

one at time $t = 2$. If the location choice is solved sequentially using our heuristic algorithm, then at time $t = 1$, the central authority will open one branch in M_3, which generates a profit of $\pi' + \pi/2 + \pi/2 = \pi + \pi'$, higher than the profit of $\pi + \pi'/2$ generated from opening a branch in either M_1 or M_2. At time $t = 2$, given that there is already a branch in M_3, the central authority would open a branch in either M_1 or M_2, which generates a profit of π by itself and at the same time reduces the profit of the branch in M_3 by $\pi/2$ due to cannibalization. Thus, the total profit of the two branches is $1.5\pi + \pi'$. Instead, if the two branch locations were solved jointly to maximize total profits, then the central authority should open the two branches in M_1 and M_2. The total profit would be $2\pi + 0.5\pi'$, which is higher than $1.5\pi + \pi'$ because $\pi' < \pi$.

Cases like this are rare because multiple conditions need to be satisfied. First, we need to have at least a cluster of three markets in which the central market is close enough to the peripheral markets, which are far away from each other. However, there are not many clusters with more than three markets in our calibrated Thai economy. Second, the central market itself needs to be small, belong to group 2, and generate less profits than peripheral markets. In other words, the profits of opening a branch in the central market are derived mainly from serving households from nearby peripheral markets, not from households living in the central market. However, it is hard to find such central markets because the loan demand declines significantly when $d_{i,t}$ increases from 0 to 10 minutes (see panel C of Figure OA.12). Third, even if we can find such clusters of central and peripheral markets, they matter for our solution only on the margin, which rarely occurs. For example, in our illustrative case above, our heuristic algorithm would return a globally optimal solution if the central authority opens only one branch at time $t = 1, 2$, or if the central authority opens two branches at time $t = 1$ and zero at time $t = 2$, or if the central authority opens two branches at time $t = 2$ and zero at time $t = 1$, or if the central authority opens more than two branches at time $t = 1, 2$. In other words, our heuristic algorithm cannot return a globally optimal solution only in the case in which the central authority opens one branch at time $t = 1$ and one at time $t = 2$.

As a robustness check, instead of solving the location choice problem every year, we solve it every 2 year (and every 3 years), which is computationally feasible and ensures that the locations of all branches opened in 2 (and 3) adjacent years are jointly chosen to maximize their total profits. We find that the solution for branch locations is exactly the same as solving the problem every year.

5.6 Evaluation of the Accuracy of the Algorithm

In this section, we evaluate the accuracy of our numerical algorithm at the province level. We cannot evaluate the accuracy at the country level because the exactly optimal location choice strategy is not attainable, given the NP-hard nature of the problem.

Table OA.15: The distribution of the province-level number of combinations.

Number of combinations	Number of provinces
< 100	16
$[100\ ,\ 1,000)$	8
$[1,000\ ,\ 10,000)$	15
$[10,000\ ,\ 100,000]$	5
$> 100,000$	32

Using the calibrated parameters and equilibrium interest rates and wages solved in our baseline specification, we solve the location choice for each province separately. The number of branches opened in each year in the province is exogenously given by the data. We apply our proposed algorithm to find the "approximately optimal" location choice for the province, which is compared with the "exactly optimal" location choice obtained using a brute-force algorithm. The brute-force algorithm produces the "exactly optimal" location choice because it exhausts all possible location choices to find the one that yields the highest profits in each year in the province.

Applying the brute-force algorithm involves solving the profits for all possible combinations of new branch locations. Even at the province level, the number of possible combinations could be huge, making the brute-force algorithm infeasible. For example, the number of combinations for some province o is calculated as follows. Consider 1986 as $t = 0$ and 1996 as $t = 10$. Denote by N_o the number of potential locations in year $t = 0$ and by $n_{o,t}$ the number of new branches opened in year t in the data. In year t, $n_{o,t}$ locations are chosen from $N_o - \sum_{\tau=1}^{t-1} n_{o,\tau}$ potential locations, giving $\binom{N_o - \sum_{\tau=1}^{t-1} n_{o,\tau}}{n_{o,t}}$ combinations. From year $t = 1$ to year $t = 10$, in total, there are in total $\Pi_{t=1}^{10} \binom{N_o - \sum_{\tau=1}^{t-1} n_{o,\tau}}{n_{o,t}}$ possible combinations by the multiplication rule. Table OA.15 summarizes the distribution of the province-level number of combinations across all provinces. It is only feasible to implement the brute-force algorithm for the provinces with fewer than $100,000$ combinations, of which there are 44.[38]

We find that for all 44 selected provinces, our algorithm gives exactly the same location choices as those produced by the brute-force algorithm. This implies that our algorithm can produce accurate results at least for the small-scale problems that we evaluate. However, we cannot evaluate the performance of the algorithm for problems of significantly larger scale due to computing limitations. Our conjecture is that even at the country level, the location choice solved using our algorithm is close to the exactly optimal solution because the partition method that we apply in Online Appendix 5.3 ensures that the interaction among markets in different segments is negligible.

[38] Our algorithm is implemented on a workstation with 56 cores (two Intel Xeon Platinum 8280 2.7GHz CPU processors, each with 28 cores). Calculating the profits of one possible combination takes approximately 3 seconds. We set the cutoff of the number of combinations to be $100,000$ such that the brute-force algorithm can be completed in about 6 hours.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105(490): 493–505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*, 59(2): 495–510.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review*, 93(1): 113–132.

Alba, Pedro, Leonardo Hernandez, and Daniela Klingebiel. 1999. "Financial Liberalization and the Capital Account: Thailand 1988-1997." The World Bank Manuscript.

Almeida, Heitor, Igor Cunha, Miguel A. Ferreira, and Felipe Restrepo. 2017. "The Real Effects of Credit Ratings: The Sovereign Ceiling Channel." *The Journal of Finance*, 72(1): 249–290.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. 2021. "Synthetic Difference in Differences."

Atanasov, Vladimir, Vladimir Ivanov, and Kate Litvak. 2012. "Does Reputation Limit Opportunistic Behavior in the VC Industry? Evidence from Litigation against VCs." *Journal of Finance*, 67(6): 2215–2246.

Azoulay, Pierre, Joshua Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *The Quarterly Journal of Economics*, 125(2): 549–589.

Balsmeier, Benjamin, Lee Fleming, and Gustavo Manso. 2017. "Independent boards and innovation." *Journal of Financial Economics*, 123(3): 536–557.

Bennedsen, Morten, Francisco Pérez-González, and Daniel Wolfenzon. 2020. "Do CEOs Matter? Evidence from Hospitalization Events." *Journal of Finance*, 75(4): 1877–1911.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Buera, Francisco J., and Yongseok Shin. 2011. "Self-insurance vs. self-financing: A welfare analysis of the persistence of shocks." *Journal of Economic Theory*, 146(3): 845–862.

Buera, Francisco J., and Yongseok Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy*, 121(2): 221 – 272.

Campello, Murillo, and Mauricio Larrain. 2016. "Enlarging the Contracting Space: Collateral Menus, Access to Credit, and Economic Activity." *Review of Financial Studies*, 29(2): 349–383.

Couch, Kenneth, and Dana W. Placzek. 2010. "Earnings Losses of Displaced Workers Revisited." *American Economic Review*, 100(1): 572–89.

Dabla-Norris, Era, Yan Ji, Robert M. Townsend, and D. Filiz Unsal. 2021. "Distinguishing Constraints on Financial Inclusion and Their Impact on GDP, TFP, and The Distribution of Income." *Journal of Monetary Economy*, 117: 1–18.

Dudzinski, Krzysztof, and Stanislaw Walukiewicz. 1987. "Exact methods for the knapsack problem and its generalizations." *European Journal of Operational Research*, 28(1): 3–21.

Eckert, Fabian, and Costas Arkolakis. 2017. "Combinatorial Discrete Choice." Working Papers 249.

Felkner, John S., and Robert M. Townsend. 2011. "The Geographic Concentration of Enterprise in Developing Countries." *The Quarterly Journal of Economics*, 126(4): 2005–2061.

Goetz, Martin R., Luc Laeven, and Ross Levine. 2016. "Does the geographic expansion of banks reduce risk?" *Journal of Financial Economics*, 120(2): 346–362.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies*, 64(4): 605–654.

Holmes, Thomas J. 2011. "The Diffusion of Wal-Mart and Economies of Density." *Econometrica*, 79(1): 253–302.

Iacus, Stefano M., Gary King, and Giuseppe Porro. 2012. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis*, 20(1): 1–24.

Jager, Simon, and Jörg Heining. 2019. "How Substitutable Are Workers? Evidence from Worker Deaths." MIT Working Paper.

Jeong, Hyeok, and Robert M. Townsend. 2008. "Growth And Inequality: Model Evaluation Based On An Estimation-Calibration Strategy." *Macroeconomic Dynamics*, 12(S2): 231–284.

Jeong, Hyeok, and Robert Townsend. 2007. "Sources of TFP growth: occupational choice and financial deepening." *Economic Theory*, 32(1): 179–221.

Jia, Panle. 2008. "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." *Econometrica*, 76(6): 1263–1316.

Ji, Yan, Songyuan Teng, and Robert M. Townsend. 2022. "Dynamic Bank Expansion: Spatial Growth, Financial Access, and Inequality." Working Paper.

Jones, Charles I., and Peter J. Klenow. 2016. "Beyond GDP? Welfare across Countries and Time." *American Economic Review*, 106(9): 2426–2457.

Kaplan, Greg, and Giovanni L. Violante. 2014. "A Model of the Consumption Response to Fiscal Stimulus Payments." *Econometrica*, 82(4): 1199–1239.

Karp, Richard M. 1972. "Reducibility among Combinatorial Problems." *Complexity of Computer Computations*, , ed. Raymond E. Miller, James W. Thatcher and Jean D. Bohlinger, 85–103. Boston, MA:Springer US.

Kaufman, Leonard, and Peter J. Rousseeuw. 2005. *Finding groups in data an introduction to cluster analysis. Wiley series in probability and statistics*, Hoboken, N.J.:Wiley-Interscience.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211–251.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *American Economic Review*, 104(10): 3186–3221.

Nelson, Andrew, Alexander de Sherbinin, and Francesca Pozzi. 2006. "Towards Development of a High Quality Public Domain Global Roads Database." *Data Science Journal*, 5: 223–265.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre Daniel Sarte, and Nicholas Trachter. 2020. "Plants in Space." National Bureau of Economic Research, Inc NBER Working Papers 27303.

Parra-Hernandez, Rafael, and Nikitas J. Dimopoulos. 2005. "A New Heuristic for Solving the Multichoice Multidimensional Knapsack Problem." *IEEE Transactions on Systems, Man, and Cybernetics*, 35(5): 708–717.

Paulson, Anna, Robert Townsend, and Alexander Karaivanov. 2006. "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship." *Journal of Political Economy*, 114(1): 100–144.

Paweenawat, Archawa, and Robert M. Townsend. 2019. "The Impact of Regional Isolationism: Disentangling Real and Financial Factors." Department of Economics, MIT Working paper.

Ren, Zhigang, Zuren Feng, and Aimin Zhang. 2012. "Fusing ant colony optimization with Lagrangian relaxation for the multiple-choice multidimensional knapsack problem." *Information Sciences*, 182: 15–29.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika*, 70(1): 41–55.

Rubin, Donald B. 1980. "Bias Reduction Using Mahalanobis-Metric Matching." *Biometrics*, 36(2): 293–298.

Rubin, Donald B. 2006. "Matching to Remove Bias in Observational Studies." *Matched Sampling for Causal Effects*, 62–80. Cambridge University Press.

Rysman, Marc, Robert M. Townsend, and Christoph Walsh. 2022. "Branch Location Strategies and Financial Service Access in Thai Banksing." Working Papers.

Samphantharak, Krislert, and Robert M. Townsend. 2018. "Risk and Return in Village Economies." *American Economic Journal: Microeconomics*, 10(1): 1–40.

Smith, Jeffrey, and Petra Todd. 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?" *Journal of Econometrics*, 125(1-2): 305–353.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick. 2019. "Capitalists in the Twenty-First Century." *The Quarterly Journal of Economics*, 134(4): 1675–1745.

Townsend, Robert M. 2011. *Financial Systems in Developing Economies: Growth, Inequality and Policy Evaluation in Thailand.* Oxford University Press.

Townsend, Robert M., Sombat Sakunthasathien, and Rob Jordan. 2013. *Chronicles from the Field: The Townsend Thai Project.* Vol. 1 of *MIT Press Books*, The MIT Press.

Yu, Tao, Yue Zhang, and Kwei-Jay Lin. 2007. "Efficient Algorithms for Web Services Selection with End-to-End QoS Constraints." *ACM Transactions on the Web*, 1(1): 1–26.

Zheng, Fanyin. 2016. "Spatial Competition and Preemptive Entry in the Discount Retail Industry."