ONLINE APPENDIX

A Bank Competition and the Equilibrium Debt Contract

In Section 2.1, we showed that under certain conditions, a franchisee's optimal effort is decreasing in D (repayment – collateral). In this section, we show that repayment is decreasing in collateral in equilibrium when banks compete in a Bertrand fashion. As a result, the difference (repayment – collateral) is decreasing in collateral. Combined with the result in Section 2.1, this implies that a franchisee's optimal effort is increasing in her collateral.

A bank's profit depends on whether the franchisee defaults or not. If she does not default, the bank gets the repayment $D+C$. (Recall that D denotes the difference between the repayment and the collateral.) If the franchisee defaults, the bank seizes her collateral, C.[1] Therefore, for given effort e, a bank's expected profit is

$$(D+C)\,[1-F(\hat{\theta}\,(e,D))]+CF(\hat{\theta}\,(e,D))-I, \tag{A.1}$$

where the default critical value of the profit shock $\hat{\theta}\,(e,D)$ is defined by the default condition (see equation (1) in the main text). Given that the function $\hat{\theta}\,(e,D)$ is strictly decreasing in e (see Section 2.1), we can define its "inverse" $e(\hat{\theta},D)$. In other words, we have

$$(1-s)\,G(\hat{\theta},e(\hat{\theta},D))-D=0. \tag{A.2}$$

Note that $e(\hat{\theta},D)$ is decreasing in $\hat{\theta}$ and increasing in D. This change of variable will be helpful to prove our results below.

At the Bertrand equilibrium, a bank's expected profit is 0. Using the above change of variables, we can write the zero-profit condition as

$$D[1-F(\hat{\theta})]-I+C=0. \tag{A.3}$$

For fixed I and C, let the solution to the above equation (A.3) be $\hat{\theta}_b\,(D)$, where the subscript b stands for "bank".

Lemma 1 *The break-even schedule $\hat{\theta}_b\,(D)$ is continuous and strictly increasing in D.*

To see this, note that the partial derivative of the bank's expected profit with respect to D is $1-F(\hat{\theta})>0$, and the partial derivative with respect to $\hat{\theta}$ is $-Df(\hat{\theta})<0$. By the implicit function theorem, the break-even schedule $\hat{\theta}_b\,(D)$ is continuously differentiable and strictly increasing in D.

[1] We assume the scrap value to be zero for simplicity. However, our results hold for any scrap value smaller than the uncollateralized portion of the repayment, D.

Intuitively, this means that for given C and I, as the total repayment increases (i.e., as $R = D + C$ increases), the bank can accept a higher default probability and still break even.

Using the same change of variable, we can also rewrite the franchisee's problem as choosing $\hat{\theta}$ with effort e adjusting according to $e(\hat{\theta}, D)$. In other words, her objective function (2) can be rewritten as follows:

$$W(\hat{\theta}, D) = U(0) F(\hat{\theta}) + \int_{\hat{\theta}}^{\infty} U\Big((1-s) G(\theta, e(\hat{\theta}, D)) - D\Big) dF(\theta) - \Psi(e(\hat{\theta}, D)). \tag{A.4}$$

Denote the solution to the above franchisee's utility maximization problem by $\hat{\theta}_f(D)$, where the subscript f stands for "franchisee". The lemma below states that the optimal default cutoff is strictly increasing in D.

Lemma 2 *$\hat{\theta}_f(D)$ is strictly increasing in D.*

We prove this lemma using standard monotone comparative statics arguments by showing that $\frac{\partial^2 W(\hat{\theta}, D)}{\partial \hat{\theta} \partial D} > 0$. See Section A.1.

We now put the two schedules $\hat{\theta}_b(D)$ and $\hat{\theta}_f(D)$ together to discuss the existence of a debt-contract equilibrium and its comparative statics.

Proposition 3 *There exists a generically unique equilibrium. In this equilibrium, either no debt contract is offered or the equilibrium debt contract has the property that the repayment is decreasing in the collateral C and increasing in I.*[2]

The proof consists of two steps. In Step 1, we show that the schedule $\hat{\theta}_f(D)$ is above the schedule $\hat{\theta}_b(D)$ at the lower bound of D. Note that the lower bound for equilibrium D is $I - C$ because $D = R - C$ and repayment R is at least weakly larger than I. Step 1 implies that the schedule $\hat{\theta}_f(D)$ either lies everywhere above the schedule $\hat{\theta}_b(D)$, or that when it crosses the schedule $\hat{\theta}_b(D)$ for the first time, it crosses from above. Intuitively, the former case where no debt contract is offered in equilibrium can occur when C is very small. In the latter case, we show in Step 2 that only the first intersection of the two schedules is an equilibrium and that the equilibrium has the property stated in the proposition.

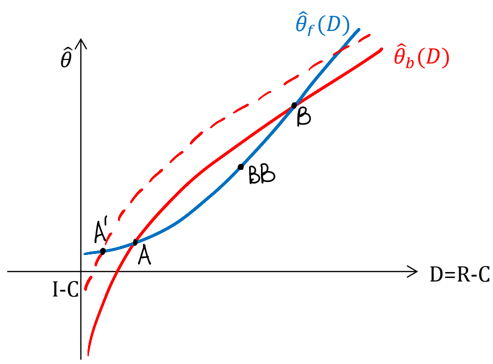
In what follows, we provide a sketch of the proof and further intuition about the two steps. Section A.2 provides the formal proof of Proposition 3.

In Step 1, we show that $\lim_{D \to I-C} \hat{\theta}_f(D) > -\infty$ and $\lim_{D \to I-C} \hat{\theta}_b(D) = -\infty$, and hence $\lim_{D \to I-C} \hat{\theta}_f(D) > \lim_{D \to I-C} \hat{\theta}_b(D)$. Intuitively, for a franchisee, even when D is as small as $I - C$, her optimal probability of default is not 0. But for a bank, when $D = I - C$, i.e., when the repayment equals I, it breaks even if and only if the franchisee never defaults. Step 1 implies

[2] In some cases, there might be multiple equilibria. As described in Section A.2, debt contracts in all equilibria have the same property as stated in Proposition 3.

that the schedule $\hat{\theta}_f(D)$ either lies everywhere above the schedule $\hat{\theta}_b(D)$, or that when it crosses the schedule $\hat{\theta}_b(D)$ for the first time, it crosses from above. Intuitively, the former case can occur when C is very small. In the latter case (see Figure A.1),[3] we can show in Step 2 that only the first intersection of the two schedules is an equilibrium (Point A in Figure A.1). This is because at any other intersection, a bank has an incentive to deviate. Take Point B (a potential equilibrium point) and Point BB (a potential deviating point) as an example. A contract at Point BB requires a smaller D than a contract at Point B so that it would be preferred by the franchisee. Given that Point BB is below the break-even schedule $\hat{\theta}_b(D)$, the deviating bank earns a positive profit. In other words, this is a profitable deviation so Point B cannot be an equilibrium. This leaves the intersection A as the unique equilibrium. Note that the schedule $\hat{\theta}_f(D)$ lies above the schedule $\hat{\theta}_b(D)$ to the left of the equilibrium point. When C increases or I decreases by 1, the schedule $\hat{\theta}_b(D)$ shifts to the left by more than 1 (as implied by equation (A.3)), and the equilibrium moves from Point A to Point A'. Given that the schedule $\hat{\theta}_f(D)$ is increasing, the equilibrium D is decreased by more than 1, implying a decrease in repayment (because repayment $R = D + C$).

Figure A.1: Equilibrium



In summary, we have shown that the equilibrium repayment is decreasing in collateral C. Combined with our result in Section 2.1 that a franchisee's optimal effort is decreasing in $D(=$ repayment $- C)$, this means that the franchisee's optimal effort and hence the franchisor's profit is increasing in C. This is further confirmed in our numerical example in Section 2.3.

[3]This figure shows a case where both schedules are continuous. We have shown in Lemma 1 that $\hat{\theta}_b(D)$ is indeed continuous; see Section A.2 for cases where the schedule $\hat{\theta}_f(D)$ has discontinuous jumps. Note that such jumps occur only when the franchisee's utility maximization problem has multiple solutions.

A.1 Proof of Lemma 2

As explained above, we need to show $\frac{\partial^2 W(\hat{\theta},D)}{\partial\hat{\theta}\partial D} > 0$ to prove this lemma. Note that

$$\frac{\partial W(\hat{\theta}, D)}{\partial \hat{\theta}} = \int_{\hat{\theta}}^{\infty} U'\left((1-s)\, G(\theta, e(\hat{\theta}, D)) - D\right)(1-s)\, G_e(\theta, e(\hat{\theta}, D))dF(\theta)\frac{\partial e}{\partial \hat{\theta}} - \Psi'(e(\hat{\theta}, D))\frac{\partial e}{\partial \hat{\theta}}, \tag{A.5}$$

and thus

$$\frac{\partial^2 W\left(\hat{\theta}, D\right)}{\partial\hat{\theta}\partial D} = \int_{\hat{\theta}}^{\infty} U''(\cdot)\left[(1-s)\, G_e(\cdot,\cdot)\frac{\partial e}{\partial D} - 1\right](1-s)\, G_e(\cdot,\cdot)\, dF(\theta)\frac{\partial e}{\partial\hat{\theta}} + \int_{\hat{\theta}}^{\infty} U'(\cdot)(1-s)\, G_{ee}(\cdot,\cdot)\, dF(\theta)\frac{\partial e}{\partial\hat{\theta}}\frac{\partial e}{\partial D} - \Psi''(\cdot)\frac{\partial e}{\partial\hat{\theta}}\frac{\partial e}{\partial D}, \tag{A.6}$$

where we omit some arguments of a function for the sake of readability. Recall that $\frac{\partial e}{\partial\hat{\theta}} < 0$ and $\frac{\partial e}{\partial D} > 0$. Since $G_{e\theta} \geq 0$ by assumption, $G_e(\theta, e(\hat{\theta}, D)) \geq G_e(\hat{\theta}, e(\hat{\theta}, D))$ for $\theta \geq \hat{\theta}$. From equation (1), we know that $(1-s)\, G_e(\hat{\theta}, e(\hat{\theta}, D))\frac{\partial e}{\partial D} = 1$. Therefore, in the first term in (A.6), $(1-s)\, G_e(\cdot,\cdot)\frac{\partial e}{\partial D} > 1$. As a result, the first term is non-negative given that $U'' < 0$ and $G_e > 0$. The second term is also non-negative under the assumption that $G_{ee} \leq 0$. Finally, the third term is positive under the assumption that $\Psi'' > 0$. Q.E.D.

A.2 Proof of Proposition 3

As mentioned, we prove Proposition 3 in two steps.

Step 1. We show that $\lim_{D\to I-C}\hat{\theta}_f(D) > -\infty$ and $\lim_{D\to I-C}\hat{\theta}_b(D) = -\infty$, and hence $\lim_{D\to I-C}\hat{\theta}_f(D) > \lim_{D\to I-C}\hat{\theta}_b(D)$.

We prove $\lim_{D\to I-C}\hat{\theta}_f(D) > -\infty$ by first noting that the franchisee's optimal effort is always bounded. Under the assumption that $G_e(\theta, e)$ is bounded and the regularity condition that $\lim_{w\to\infty} U'(w) < \infty$ (which implies that U' is bounded), the first term in the first-order condition (3) is bounded. However, $\lim_{e\to\infty}\Psi'(e) = \infty$. Therefore, the first-order condition (3) implies that the optimal effort is bounded. Given that $\lim_{\theta\to-\infty} G(\theta, e) = -\infty$ for any finite e, the left-hand side of equation (A.2) approaches $-\infty$ if $\lim_{D\to I-C}\hat{\theta}_f(D) = -\infty$. This is a contradiction to equation (A.2).

We prove $\lim_{D\to I-C}\hat{\theta}_b(D) = -\infty$ also by contradiction. Suppose $\lim_{D\to I-C}\hat{\theta}_b(D) > -\infty$.

Then, when we take the limit of the left-hand side of equation (A.3), we have

$$\begin{aligned}
& \lim_{D \to I-C} D\left[1 - F(\hat{\theta}_b(D))\right] + \int_{-\infty}^{\hat{\theta}_b(D)} (1-s)\, G\left(\theta, e\left(\hat{\theta}_b(D), D\right)\right) dF(\theta) - I + C \qquad \text{(A.7)} \\
< \;& \lim_{D \to I-C} D\left[1 - F(\hat{\theta}_b(D))\right] + \int_{-\infty}^{\hat{\theta}_b(D)} (1-s)\, G\left(\hat{\theta}_b(D), e\left(\hat{\theta}_b(D), D\right)\right) dF(\theta) - I + C \\
= \;& \lim_{D \to I-C} D\left[1 - F(\hat{\theta}_b(D))\right] + (1-s)\, G\left(\hat{\theta}_b(D), e\left(\hat{\theta}_b(D), D\right)\right) F(\hat{\theta}_b(D)) - I + C \\
= \;& \lim_{D \to I-C} D - I + C = 0
\end{aligned}$$

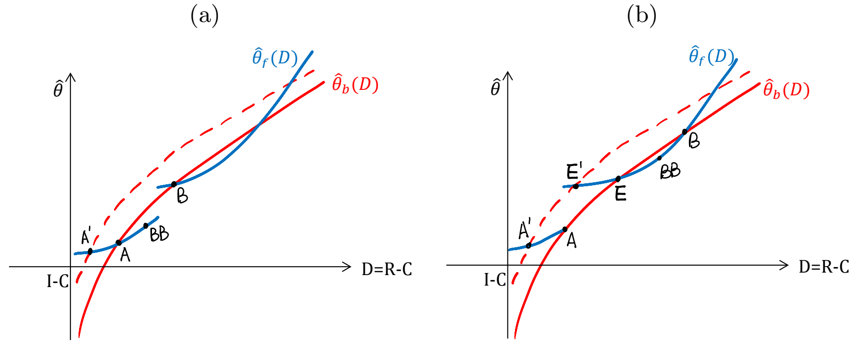
which is a contradiction to equation (A.3). In (A.7), the first inequality holds because $G(\theta, e)$ is increasing in θ. The last equality holds because of equation (A.2).

Step 2. We put the two schedules $\hat{\theta}_b(D)$ and $\hat{\theta}_f(D)$ together to discuss the equilibrium and the comparative statics.

Step 1 implies that the schedule $\hat{\theta}_f(D)$ either lies everywhere above the schedule $\hat{\theta}_b(D)$, or that when it crosses the schedule $\hat{\theta}_b(D)$ for the first time, it crosses from above. Figure A.1 shows a case where the two schedules intersect twice. We have shown that the only equilibrium is Point A in Figure A.1 and that the equilibrium repayment is decreasing in C.

Figure A.1 shows a case where both schedules are continuous. We have shown in Lemma 1 that $\hat{\theta}_b(D)$ is indeed continuous. The schedule $\hat{\theta}_f(D)$, however, can have discontinuous jumps. Such jumps occur only when the franchisee's utility maximization problem has multiple solutions. That is why we consider such a case non-generic. Given that $\hat{\theta}_f(D)$ is increasing in D as shown in Lemma 2, all jumps are upward jumps. Figures A.2(a) and A.2(b) show two possible cases with discontinuous $\hat{\theta}_f(D)$ where the two schedules intersect more than once (if the two schedules intersect only once, the uniqueness is established automatically). In Figure A.2(a), the only equilibrium is Point A, as Point BB would be a profitable deviation for any other intersection of the two schedules. As a result, the comparative statics result holds. In Figure A.2(b), both Point A and Point E are equilibria. But the comparative statics result still holds: as C increases or I decrease by 1, the schedule $\hat{\theta}_b(D)$ shifts to the left by more than 1, so the equilibrium repayment R decreases. Q.E.D.

Figure A.2: Equilibrium



B Details on the Numerical Example in Subsection 2.3

In this section, we describe the numerical example for our analysis in Subsection 2.3. In this numerical example, the amount of capital needed to open an outlet is $I = 93,000$, which is the average capital needed in the data in 82-84 dollars. We assume that the franchisor's and the franchisee's decisions are based on their calculation over a ten-year horizon, after which we assume that the capital is worth nothing.

We assume a CARA utility function for the franchisee: $U(w) = -\exp(-\rho w)$ where ρ is her parameter of absolute risk aversion and w is her payoff. We use the median absolute risk aversion parameter estimated by Cohen and Einav (2007): $\rho = 3.4E-5$. We assume that her cost of effort is $\Psi(e) = \kappa \exp(1/(1-e))$, where e is measured as a fraction of available working hours. Note that the cost of effort is infinity if $e = 1$.[4] We choose $\kappa = 0.034$ (see below) and consider that there are 140 available working hours per week.[5]

We normalize a hired manager's effort to be $e_0 = 40/140$. The expected profit from a company-owned outlet is $\tilde{\pi}_c = E_\theta [G(\theta, e_0)] - I - W$, where $W = 200,000$ is the total wage for a hired manager over ten years. An annual wage of 20,000 in 82-84 dollars is consistent with the wage for a typical manager working for a chain.

We assume a linear functional form for $G(\theta, e)$: $G(\theta, e) = \theta + \lambda e$, where $\lambda = 70,000$ (see below) and the profit shock θ follows a normal distribution with mean $\bar{\theta}$ and variance σ_θ^2. We choose $\bar{\theta} = 325,000$ so that $\bar{\theta} + \lambda e_0$ at $e_0 = 40/140$ (a hired manager's effort) is 345,000, which amounts to an annual return on investment of around 15%.[6] We assume that $\sigma_\theta = \bar{\theta}$ so that earning zero profit is one standard deviation away from the mean.

As for the franchise contract, we assume that the royalty rate is 10%. According to Blair and Lafontaine (2005), the mode royalty rate is 5% of revenue. Since the function $G(\theta, e) = \theta + \lambda e$ is in fact the present value of the *profit* flow, we use 10%. A franchise contract typically also specifies a one-time lump-sum franchise fee that a franchisee pays a franchisor. Given that it is a lump-sum fixed cost for the franchisee (i.e., it is a decision-irrelevant constant component in her payoff), we omit it in the franchisee's problem in the theoretical model. But it does add to the profit for the franchisor. We assume that the franchise fee is $L = 13,500$, which is consistent with the mean of the franchise fee in the data (in 82-84 dollars).

We have little guidance on the parameters κ and λ. We choose $\kappa = 0.034$ and $\lambda = 70,000$ so that (1) the marginal cost of effort at e_0 is consistent with the hourly wage of a hired manager;[7]

[4] In Section 2, we impose that $\lim_{e\to\infty} \Psi(e) = \infty$. There, the notation "$e$" is equivalent to $1/(1-e)$ here. We use this change of variable in this calibration exercise so that effort has a more concrete meaning.

[5] Bitler, Moskowitz and Vissing-Jorgensen (2005) find that the median entrepreneur works for 48 hours per week in their sample. They consider this amount of working hours as equivalent to working 1/3 of available hours, implying that an entrepreneur has 144 "available" hours for work per week.

[6] Ignoring discounting for simplicity, this amount means an annual profit of 34,500, 20,000 of which is compensating the franchisee's time, leaving 14,500 as the return to an investment of 93,000.

[7] An additional hour per week increases "effort" e by 1/140. The marginal cost of effort is therefore $\Psi'(e_0)/140$,

and (2) a franchisee's optimal effort is in a reasonable range, i.e., between 15 and 80 hours per week depending on the amount of collateral.

Opening an outlet in this chain requires capital I. In the debt contract, the repayment depends on the amount of money borrowed (I) and the collateral (C) according to the following linear function: repayment $= (1+r)\,I$ where $r = 0.35 - (0.35 - 0.01)C/I$. In other words, the interest rate is 35% (over the life of a loan) when $C = 0$ and 1% when $C = I$. We have also done the calibration exercise assuming a constant interest rate of 25% (i.e., under the assumption that the repayment is 1.25 times I), which generates very similar results.

We assume that there are $N = 10$ potential entrants. The collateralizable wealth that each potential franchisee has follows a log-normal distribution with mean $\bar{C}$, which we allow to vary to generate the results of interest, and a standard deviation equal to $34,000$, which is close to the standard deviation of the average collateralizable wealth in a state across states in the data (33,400).

C Data Appendix

This section provides further details on data and measurement issues.

C.1 Franchisor Sample and Characteristics

We constructed our sample of franchised chains from yearly issues of the Entrepreneur Magazine from 1981 to 1993, and an annual listing called the *Bond Franchise Guide* (previously the *Source Book of Franchise Opportunities*) from 1994 to 2007. In each case, the publication is a year late relative to the year of data collection, so we obtain the 1980 to 1992 data from the first source and the 1993 to 2006 data from the second. Since the *Bond Franchise Guide* was not published in 2000 and 2003, we are missing data for all franchisors for 1999 and 2002.

Because data on collateralizable housing wealth are only available from 1984 onward, we constrain our sample to U.S.-based franchisors that started their business in 1984 or later. This means that our sample comprises mostly young brands, with a small number of establishments: well-known brands such as McDonald's and Burger King, for example, were established in the 1950s and 1960s and are absent from our data. Our data sources provide information on 1016 U.S.-based franchisors started their business in 1984 or later.

After eliminating hotel chains (for reasons given in footnote 17), and deleting observations for outlier franchisors who either grow very fast (the number of outlets increases by more than 100 in

implying an annual marginal cost of $\Psi'(e_0)/1400$. For an hourly wage of $\frac{20,000}{250\times 8}$ (=10, considering 250 working days in a year and 8 working hours in a day), an additional hour per week generates a marginal utility of $10\rho\exp(-\rho w)$, where $w = 20,000$ is the annual wage. With the chosen parameters, the marginal cost of effort is close to this marginal utility.

a year) or shrink very fast,[8] our final sample consists of 3820 observations regarding 934 distinct franchised chains, with an average of four observations per chain. This low number of observations per chain is explained in part by the large amount of entry into and exit from franchising (or business) of the chains as well as the lack of data for 1999 and 2002.[9]

For each franchisor/year in our sample, we have data on the amount of capital required to open an outlet (Capital Required) and the number of employees that the typical outlet needs (Number of Employees). We transform the former to constant 1982-84 dollars using Consumer Price Index data from the Bureau of Labor Statistics. For the latter, we count part-time employees as equivalent to 0.5 of a full-time employee.

We view the Capital Required (in constant dollars) and the Number of Employees needed to run the business as intrinsically determined by the nature of the business concept, which itself is intrinsically connected to the brand name. So, they should not change from year to year. Yet we find some variation in the data. Since the data are collected via surveys, they are subject to some errors from respondents or transcription. We therefore use the average across all the observations we have for these two variables for each franchised chain under the presumption that most of the differences over time reflect noise in the survey data collected by our sources. There is also some variation in the reported years in which the chain begins franchising and when it starts its business. For these variables, we use the earliest date given because we see that franchisors sometimes revise these dates to more current values for reasons we do not fully understand. However, we do make sure that the year of first franchising is after the first year in business. We also push the year of franchising to later if we have data indicating no franchised establishments in the years when the chain states it started franchising.

C.2 Collateralizable Housing Wealth

We measure collateralizable housing wealth using

- data on a yearly housing price index at the state level from the Federal Housing Finance Agency. These data are revised at the source quite frequently, perhaps as often as every time a new quarter is added. They also have been moved around several web sites. The version used here is the "States through 2010Q3 (Not Seasonally Adjusted) [TXT/CSV]" series in

[8] A franchisor is considered shrinking very fast if more than half of the existing outlets exit in a year and the probability of such an amount of exits is less than 1e-10 assuming the exit rate of each outlet is as high as 50%. We impose the second criterion to avoid removing small chains for which a halving of the number of outlets is not an unlikely event. To give a concrete example: we consider a decrease from, say, 88 outlets to 1 outlet as shrinking very fast because the probability of this change is 2.8e-25 under the assumption of a 50% exit rate for each outlet ($B(87, 88, 0.5) =$ 2.8e-25, where B is the binomial probability function). Given such a small probability, we consider such events as outliers. This is a concrete example of one of the two chains we consider as shrinking very fast and therefore drop from the sample. On the other hand, the probability of a change from 3 outlets to 1 is $B(2, 3, 0.5) = 0.38$, that is, a rather ordinary event.

[9] See e.g. Blair and Lafontaine (2005) for more on the entry and exit rate of chains.

the All-Transactions Indexes section at http://www.fhfa.gov/Default.aspx?Page=87. The base period of the index is 1980Q1;

- data on housing values by state in 1980 from the Census Bureau (the base year of the aforementioned housing price index). These data are in constant year 2000 based dollars. We transform them to constant 1983-84 based constant dollars using the Consumer Price Index.

The combination of the above two sets of data allows us to generate time series of yearly housing values per state, from 1980 onward. We then complement these with the following:

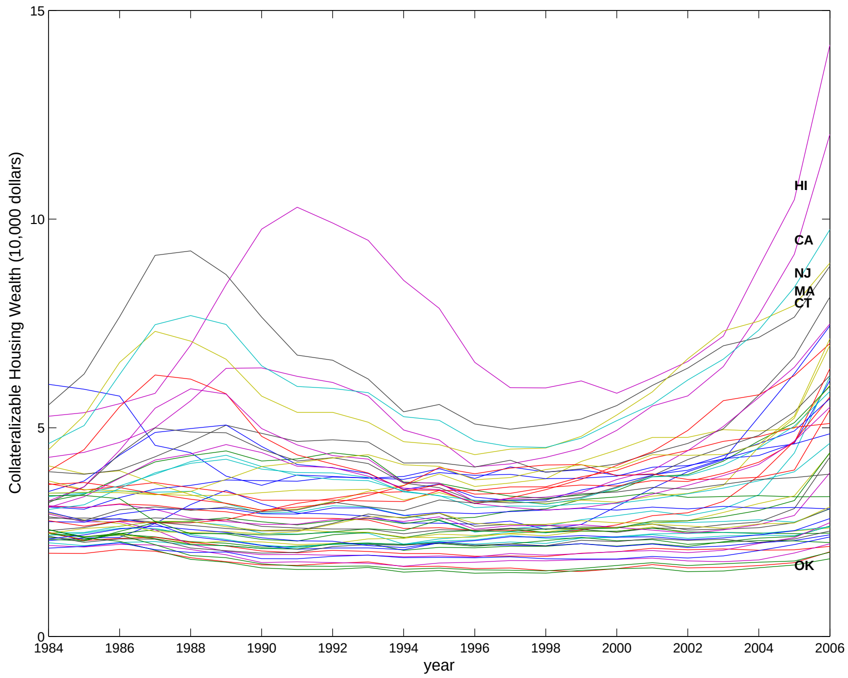
- yearly data about home ownership rates across states from the Census Bureau's Housing and Household Economic Statistics Division;

- data from the joint Census-Housing and Urban Development (HUD) biennial reports, based on the American Housing Surveys, which summarize information on mortgages on a regional basis (Northeast, Midwest, South and West). Specifically, from this source, we obtained measures of regional housing values, total outstanding principal amount, and number of houses owned free and clear of any mortgage (Tables 3-14 and 3-15 of the biennial reports). The data for housing values and for total outstanding principal are reported in the form of frequencies for ranges of values. We use the middle value for each range and the frequencies to calculate expected values for these. We then combine these data to calculate the average proportion of mortgage outstanding for homeowners in the region for each year. Specifically, we calculate $\frac{\frac{(TOPA*NTOPA)}{(NTOPA+NF)}}{(HousingValues)}$, where TOPA is Total Outstanding Principal Amount, NTOPA is the number of households that reported Total Outstanding Principal Amount, and NF is the number of households with houses owned free and clear of any mortgage. Since the data on TOPA, NTOPA, and NF are by region, we ascribe the regional expected value to all states in the region.[10] Also, since the joint Census-Housing and Urban Development (HUD) reports are biennial, we ascribe the value to the year of, and to the year before, the report. This means that we can generate our main explanatory variable of interest below from 1984 onward.

In the end, we combine the information on the proportion of outstanding mortgage for homeowners (data in the fourth item above) with the state home ownership rate (the third item) and housing value time series (a combination of the first and the second items) to calculate our measure of Collateralizable Housing Wealth for each state/year, given by: (1− the average proportion of mortgage still owed) ×(the home ownership rate)×(housing value). Figure C.1 plots collateralizable housing wealth for each state over time.

[10] We investigated several other data sources for home equity and housing values, some of which provide data at a more disaggregated level. However, none of them allowed us to go back in time as far as 1984, as our current sources do. Moreover, these sources most often covered a number of major cities but did not provide state-level data.

Figure C.1: Plot of Collateralizable Housing Wealth



C.3 Other Aggregate Economic Variables

Real Gross State Product (GSP) data are from the Bureau of Economic Analysis. We deflate nominal annual GSP data using the Consumer Price Index also from the Bureau of Labor Statistics, and obtain per capita GSP after dividing by a state's population. The annual population data are from the Census Bureau. The interest rate data series we use is the effective Federal Funds rate annual data (downloaded from the Federal Reserve web site, at www.federalreserve.gov/releases/h15/data.htm on 03/26/2009), in percent. As mentioned in the main body of the paper, the credit supply measures come from the National Foundation of Independent Business' Small Business Economic Survey and the Federal Reserve's Senior Loan Officer Opinion Survey on Bank Lending Practices.

C.4 Weighting Matrices

As described in the body of the paper, we create our main weighting matrix using information from the 1049 franchisors in our data that we observe at least once within 15 years after they start franchising. We use only one year of data per franchisor, namely the latest year within this 15 year period, to construct the matrix. For each state pair (s_1, s_2), the weight is defined as $\sum_{j \in J_{s_1}} \mathbb{1}\left(s_2 \text{ is the top state for chain } j\right) / \#\left(J_{s_1}\right)$, where J_{s_1} is the set of chains that are headquartered in state s_1, $\#(J_{s_1})$ is the cardinality of the set J_{s_1}, and $\mathbb{1}\left(s_2 \text{ is the top state for chain } j\right)$ is a dummy variable capturing whether chain j reports s_2 as the state where they have the most outlets. In other words, the weight is the proportion of chains headquartered in s_1 that report s_2

as the state where they have the most outlets. The resulting matrix is shown below as Matrix A.

We use an alternative set of weights in our robustness analysis. Our data source identifies three (or two, or one if there are only two or one) U.S. states where the chain has the most outlets, and for each of those, it indicates how many outlets it has. Our alternative weighting matrix takes all these into account, namely it uses data from all top three states (as opposed to only the top state in Matrix A) as well as the relative importance of these top three states, in the form of the proportion of outlets in each state relative to the total in all three (as opposed to only using a dummy to capture whether a state is the top state as in Matrix A). Specifically, for each chain j, we calculate $N_j = n_{1j} + n_{2j} + n_{3j}$, where n_{kj} is the number of establishments of the chain in its top three states $k = 1, 2$ or 3. We then calculate $p_{kj} = n_{kj}/N_j$. For each state pair (s_1, s_2), we calculate the average proportion of establishments in origin state s_1 and destination state s_2 pair across all the chains headquartered in state s_1 as $\sum_{j \in J_{s_1}} [p_{1j} \mathbb{1}(s_2$ is franchisor j's state with the most outlets$) + p_{2j} \mathbb{1}(s_2$ is franchisor j's state with the second most outlets$) + p_{3j} \mathbb{1}(s_2$ is franchisor j's state with the third most outlets$)]/\#\,(J_{s_1})$. Note that the sum of these average proportions across destination states s_2 for each origin state s_1 is again 1.

The resulting matrix is shown below as Matrix B. As can be seen from a comparison of the matrices, the matrix we rely on in our main specification (Matrix A) allocates some weight to macro conditions outside of the chain's headquarters state, but not as much as Matrix B does. The latter is a little more dispersed. Overall, the two matrices are similar. Consequently, the descriptive statistics in the second and third panel of Table C.1, which summarize aggregate economic variables using the two weighting schemes, are similar. Compared to using the aggregate economic variables of the home state only, with no weights, the mean and standard deviation of population in particular is quite different once we apply our weights. It is therefore important that we use these weighting matrices as this allows variation in the economic conditions of other relevant states to affect the decisions of chains headquartered in typically smaller, lower collateralizable housing wealth states.

Table C.1: Summary Statistics for Aggregate Economic Variables for Different Weighting Matrices: At the state/year level, for 48 states between 1984 and 2006

	Mean	Median	S.D.	Min	Max	Obs
No Weights						
Coll. Housing Wealth (82-84 $10K)	3.41	3.03	1.52	1.51	14.17	1104
Population (Million)	5.46	3.83	5.83	0.52	36.12	1104
Per-Capita Gross State Product (82-84 $10K)	1.85	1.73	0.67	1.09	7.47	1104
Main Matrix (Matrix A)						
Coll. Housing Wealth (82-84 $10K)	3.62	3.34	1.31	1.83	14.17	1104
Population (Million)	8.84	8.23	5.52	0.52	31.68	1104
Per-Capita Gross State Product (82-84 $10K)	1.89	1.79	0.63	1.22	7.47	1104
Alternative Matrix (Matrix B)						
Coll. Housing Wealth (82-84 $10K)	3.61	3.28	1.17	2.11	13.21	1104
Population (Million)	8.67	8.20	4.73	1.14	28.92	1104
Per-Capita Gross State Product (82-84 $10K)	1.89	1.80	0.54	1.26	6.60	1104

Matrix A: The Main Weighting Matrix Used in Constructing Values for Aggregate Economic Variables Relevant to each Chain

	AK	AL	AR	AZ	CA	CO	CT	DC	DE	FL	GA	HI	IA	ID	IL	IN	KS	KY	LA	MA	MD	ME	MI	MN	MO	MS	MT	NC	ND	NE	NH	NJ	NM	NV	NY	OH	OK	OR	PA	RI	SC	SD	TN	TX	UT	VA	VT	WA	WI	WV	WY	sum
AL	0	0.63	0	0	0	0	0	0	0	0.13	0	0	0	0	0	0	0	0	0	0	0	0	0	0.13	0	0	0	0	0	0	0	0	0	0	0.13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
AR	0.09	0	0.45	0	0	0.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0.27	0	0	0	0	0	0	0	1
AZ	0	0	0	0.42	0.35	0	0	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.06	0	0	0	0	0	0	0	0.03	0.03	0	0	0	0	0.03	0	0	1
CA	0	0	0	0.01	0.82	0.02	0	0	0	0.02	0.01	0	0	0	0	0	0	0.01	0	0	0	0	0.02	0.01	0.01	0	0	0.01	0	0	0	0	0	0.01	0.01	0.01	0	0.01	0	0	0	0	0.01	0.03	0	0	0	0.01	0	0	0	1
CO	0	0	0	0	0.17	0.65	0	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0.02	0.02	0.06	0	0	0	0	0	0	0	1
CT	0	0	0	0	0	0	0.71	0	0	0	0.05	0	0	0	0	0	0	0	0	0.10	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
DC	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
DE	0	0	0	0	0	0	0	0	0.50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.50	0	0	0	0	0	0	0	0	0	0	0	0	1
FL	0	0	0	0	0.14	0.01	0.01	0.01	0	0.62	0.03	0	0	0	0	0	0	0.01	0	0.01	0	0	0	0.01	0	0	0	0.01	0	0	0	0.02	0	0	0.04	0.01	0	0.01	0.01	0	0	0	0	0.06	0	0	0	0	0	0	0	1
GA	0	0.02	0	0	0.10	0	0	0	0	0.18	0.38	0.02	0	0	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0.07	0	0	0	0.02	0	0	0.02	0.02	0	0.02	0.02	0	0.03	0	0.02	0.08	0	0	0	0	0	0	0	1
HI	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
IA	0	0	0	0	0	0	0	0	0	0.11	0	0	0.78	0	0.11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ID	0	0	0	0	0	0	0	0	0	0	0	0	0	0.75	0	0	0	0	0	0	0	0	0	0	0	0	0.25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
IL	0	0	0	0	0.11	0	0	0	0	0.02	0	0	0	0	0.64	0.02	0	0	0	0	0	0	0.05	0.02	0	0	0	0	0	0	0	0.02	0	0	0	0.07	0	0	0	0	0	0	0	0.05	0	0	0	0	0	0	0	1
IN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0.55	0	0.09	0	0	0	0	0.09	0	0	0	0	0	0	0	0	0	0	0	0	0.18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
KS	0	0	0	0	0	0	0	0	0	0.07	0	0	0.07	0	0	0	0.33	0	0	0	0	0	0	0.07	0.13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.07	0.20	0	0	0	0	0.07	0	0	1
KY	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0.06	0	0.61	0	0	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0	0	0.06	0	0	0	0	0	0	0	0.06	0	0	0	0	0.06	0	0	1
LA	0	0	0.06	0	0.06	0	0	0	0	0	0	0	0	0.06	0	0	0	0	0.56	0	0	0	0	0	0	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.13	0	0.06	0	0	0	0	0	1
MA	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.74	0	0	0	0	0	0	0	0.04	0	0	0	0	0	0	0.04	0	0	0	0	0.04	0	0	0	0.09	0	0	0	0	0	0	0	1
MD	0	0	0	0	0.17	0	0	0	0	0.10	0	0	0	0	0	0	0.03	0	0	0	0.47	0	0.03	0	0	0	0	0	0	0	0	0	0	0	0.03	0	0	0	0.03	0	0	0	0	0.03	0	0.07	0	0	0.03	0	0	1
ME	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MI	0	0	0	0	0.08	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0.02	0	0	0.80	0	0	0	0	0	0	0	0	0.02	0	0	0	0.02	0	0	0	0	0	0	0	0.02	0	0	0	0	0	0	0	1
MN	0	0	0	0.04	0.08	0	0.02	0	0	0.04	0	0	0.02	0	0.02	0	0	0	0	0	0	0	0	0.54	0	0	0	0.02	0	0.04	0	0	0	0	0	0.10	0	0.02	0.02	0	0	0	0	0	0	0	0	0	0.04	0	0	1
MO	0	0	0	0	0	0	0	0	0	0.10	0	0	0.05	0	0.05	0	0.10	0	0	0	0	0	0	0	0.52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0.14	0	0	0	0	0	0	0	1
MS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.33	0	0	0	1
MT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NC	0	0	0	0	0.05	0	0	0	0	0.11	0.11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.74	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
ND	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.50	0	0	0	0	0.50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NE	0	0	0	0	0.14	0	0	0	0	0.03	0	0	0.03	0	0.03	0	0.03	0	0	0	0	0	0.03	0	0	0	0	0	0	0.52	0	0	0	0	0	0.03	0	0	0	0	0	0	0	0.07	0	0.03	0	0	0.03	0	0	1
NH	0	0	0	0	0	0	0	0	0	0.33	0	0	0	0	0	0	0	0	0	0.33	0	0	0	0	0	0	0	0	0	0	0.33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
NJ	0	0	0	0	0.09	0	0	0	0	0.07	0	0	0	0	0	0	0	0	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0.55	0	0	0.09	0.02	0	0	0.09	0	0	0	0.02	0.02	0	0.02	0	0	0	0	0	1
NV	0	0	0	0	0.21	0.07	0	0	0	0.07	0	0.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.14	0	0	0	0	0	0.14	0	0	0	0	0	0	0	0	0	0	0	0.07	0	0.21	0	0	0	1
NY	0	0	0	0	0.06	0	0.04	0	0	0.08	0.04	0.02	0	0.02	0	0	0	0.02	0	0.02	0	0	0.02	0	0	0	0	0	0	0	0	0.04	0	0	0.59	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0	0	1
OH	0	0	0	0	0.11	0	0	0	0	0.05	0	0	0	0	0	0	0	0	0	0	0.03	0	0.03	0	0	0	0	0.02	0	0	0	0.02	0	0	0	0.68	0	0	0.02	0	0	0	0.02	0.03	0.02	0	0	0	0	0	0	1
OK	0	0	0	0	0	0	0	0	0	0.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.25	0	0.08	0	0	0	0.08	0.50	0	0	0	0	0	0	0	1
OR	0	0	0	0.05	0.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.62	0	0	0	0	0	0	0	0	0	0.24	0	0	0	1
PA	0	0	0	0	0.02	0	0	0	0.02	0.05	0.02	0	0	0	0.02	0	0	0.02	0	0	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0.07	0	0	0	0.63	0	0.02	0	0	0.05	0	0	0	0	0	0.02	0	1
RI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.75	0	0	0	0	0	0	0	0	0	0	0	1
SC	0	0	0	0	0.11	0	0	0	0	0.11	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0.11	0	0	0	0	0.44	0	0	0	0	0	0	0	0	0	0	1
SD	0	0	0	0	0	0	0	0	0	0.10	0	0	0.10	0	0	0	0	0	0	0	0	0	0	0.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.70	0	0	0	0	0	0	0	0	0	1
TN	0	0.04	0.04	0	0.11	0	0	0	0	0.04	0.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.50	0.04	0	0.04	0	0	0	0	0	1
TX	0	0	0	0.02	0.18	0	0	0	0	0.02	0.02	0	0	0	0.01	0	0	0	0	0	0	0	0.01	0	0	0	0	0	0	0	0	0	0.02	0.01	0.01	0	0	0	0	0	0	0	0.01	0.68	0	0.01	0	0	0	0	0	1
UT	0	0	0	0	0.52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0	0	0	0	0	0.04	0.35	0	0	0	0	0	0	1
VA	0	0	0	0	0.13	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.04	0	0	0	0	0	0	0.04	0	0	0	0	0	0	0.09	0.04	0	0	0.09	0	0	0	0	0	0	0.52	0	0	0	0	0	1
VT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00	0	0	0	0	1
WA	0	0	0	0	0.06	0	0	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.12	0	0	0	0	0	0.06	0	0	0	0.71	0	0	0	1
WI	0	0	0	0	0.08	0	0	0	0	0.08	0	0	0	0	0.08	0	0	0	0	0	0	0	0.08	0.04	0	0	0	0	0	0	0	0	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0.58	0	0	1
WV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.20	0	0	0.20	0	0	0	0	0	0	0	0	0	0	0.40	0	1

For a chain whose home state is AL, for example, its aggregate economic variable is the weighted sum of the variable in the 50 states (plus DC) where the weights are given by the first row.
Weights larger than or equal to 0.10 are highlighted.

Matrix B: Alternative Weighting Matrix Used in Constructing Values for Aggregate Economic Variables Relevant to each Chain

	AK	AL	AR	AZ	CA	CO	CT	DC	DE	FL	GA	HI	IA	ID	IL	IN	KS	KY	LA	MA	MD	ME	MI	MN	MO	MS	MT	NC	ND	NE	NH	NJ	NM	NV	NY	OH	OK	OR	PA	RI	SC	SD	TN	TX	UT	VA	VT	WA	WI	WV	WY	sum
AL	0	0.35	0	0	0	0	0	0	0	0.14	0.05	0	0	0	0	0	0	0	0.02	0	0	0	0	0.06	0	0	0	0.01	0	0	0	0	0	0	0.05	0.05	0	0	0	0	0.00	0	0.19	0	0	0.05	0	0	0.01	0	0	1
AR	0.02	0	0.33	0.04	0.05	0.08	0	0	0	0.01	0	0	0	0	0	0.02	0.00	0	0.01	0	0	0	0	0	0.04	0	0	0	0	0.01	0	0	0	0	0	0	0.01	0	0	0	0	0	0.09	0.28	0	0	0	0	0	0	0	1
AZ	0	0	0	0.38	0.20	0.03	0	0.01	0	0.06	0	0	0	0	0	0	0	0	0	0	0.00	0.00	0	0	0	0	0	0.00	0	0.01	0	0.01	0	0.00	0.08	0.01	0	0.00	0.00	0	0	0	0.02	0.07	0.02	0.01	0	0.01	0.04	0	0	1
CA	0	0.00	0	0.02	0.71	0.02	0	0	0	0.02	0.02	0.00	0	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0	0.00	0.01	0.00	0.00	0	0.00	0	0.00	0.00	0.01	0.00	0.01	0.02	0.01	0.00	0.01	0.01	0	0	0	0.00	0.04	0.00	0	0	0.02	0.00	0	0	1
CO	0	0	0	0.03	0.11	0.49	0	0	0	0.12	0.01	0.00	0.01	0.02	0.01	0	0.00	0	0	0	0	0	0.00	0.00	0.04	0	0.00	0.01	0	0	0.00	0	0	0.00	0.00	0.01	0	0	0.00	0	0.00	0.01	0.01	0.08	0.01	0.00	0	0.01	0.00	0	0	1
CT	0	0	0	0	0	0.01	0.63	0	0.01	0.02	0.02	0	0	0	0	0	0	0	0	0.10	0.01	0	0.01	0	0	0	0	0	0	0	0.01	0.06	0.00	0	0.08	0.00	0	0	0.01	0	0.01	0	0	0.01	0	0.00	0	0	0	0.01	0	1
DC	0	0	0	0	0	0	0	0.83	0	0	0	0	0	0	0.17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
DE	0	0	0	0	0	0	0	0	0.37	0	0	0	0	0	0	0	0	0	0	0	0.19	0	0	0	0	0	0	0	0	0	0	0.08	0	0	0	0	0	0	0.36	0	0	0	0	0	0	0	0	0	0	0	0	1
FL	0	0.01	0.00	0.00	0.09	0.01	0.00	0.01	0	0.52	0.04	0	0	0	0.02	0	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0	0.02	0	0.00	0	0.03	0	0.00	0.04	0.02	0	0.01	0.02	0.00	0.01	0	0.00	0.08	0.00	0.01	0	0.01	0.00	0.00	0	1
GA	0	0.04	0.00	0	0.08	0.01	0	0	0	0.12	0.32	0.03	0	0	0.02	0	0	0.00	0.01	0.00	0.01	0	0.01	0.00	0	0	0	0.07	0	0	0	0.01	0	0	0.03	0.03	0.00	0.02	0.02	0	0.02	0	0.03	0.10	0	0.01	0	0	0.00	0	0	1
HI	0	0.08	0	0	0	0	0	0	0	0	0	0.92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
IA	0	0	0	0	0.01	0.01	0	0	0	0.04	0	0	0.49	0	0.19	0.01	0	0	0	0	0	0	0.07	0.08	0.02	0	0	0	0	0.01	0	0	0	0	0	0.01	0	0	0	0	0	0.02	0	0	0.02	0	0	0	0.02	0	0	1
ID	0	0	0	0.01	0	0.04	0	0	0	0	0	0	0	0.38	0	0	0.02	0	0	0	0	0	0	0	0	0	0.33	0	0.01	0.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0	0	0.09	0	0	0	1
IL	0.00	0.00	0	0.00	0.07	0	0	0	0	0.06	0.00	0	0.01	0	0.50	0.03	0	0	0	0.01	0	0	0.08	0.01	0.00	0	0	0	0	0	0	0.02	0	0	0	0.06	0.01	0	0.03	0	0.00	0	0.01	0.04	0	0.00	0	0.01	0.05	0	0	1
IN	0	0	0	0	0	0	0	0	0	0.01	0	0	0	0	0.09	0.56	0	0.13	0	0.01	0	0	0.06	0.05	0	0	0	0	0	0	0	0	0	0	0.01	0.08	0	0	0	0	0	0	0	0.00	0	0	0	0	0	0	0	1
KS	0	0	0	0.01	0.03	0.01	0	0	0	0.02	0.01	0	0.02	0	0.05	0	0.29	0	0	0	0	0	0	0.04	0.21	0	0	0.00	0	0.01	0	0	0	0.00	0.01	0.01	0.03	0	0	0	0	0	0.03	0.23	0	0	0	0	0.01	0	0	1
KY	0	0	0	0	0.05	0.01	0.03	0	0	0.11	0	0	0	0	0.02	0.09	0	0.47	0	0	0.02	0	0.03	0	0	0	0	0	0	0	0	0.01	0	0	0.03	0.02	0	0	0	0	0	0	0.07	0.01	0	0	0	0	0.03	0	0	1
LA	0	0.04	0.07	0	0.05	0	0	0	0	0.05	0	0	0	0.01	0.01	0	0	0	0.43	0.01	0	0	0	0	0	0.00	0	0.08	0	0	0	0.01	0	0	0	0	0	0	0.01	0	0.03	0	0.01	0.08	0	0.08	0	0.01	0	0	0	1
MA	0	0	0	0	0.02	0	0.05	0	0	0.04	0	0	0	0	0.01	0.00	0	0	0	0.59	0.01	0	0	0	0	0	0	0.01	0	0	0.03	0.04	0	0	0.06	0.03	0	0	0	0.03	0	0	0	0.08	0	0	0	0	0	0	0	1
MD	0	0.00	0	0	0.05	0.01	0	0.01	0	0.03	0.03	0	0	0	0.01	0	0.02	0	0	0.01	0.39	0	0.01	0	0.01	0	0	0	0	0	0	0.05	0	0.00	0.01	0.01	0	0.00	0.10	0	0	0	0	0.05	0.00	0.15	0	0	0.04	0	0	1
ME	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.92	0	0	0	0	0	0	0	0	0.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
MI	0	0.01	0	0	0.06	0.00	0	0	0	0.05	0	0	0	0	0.01	0.01	0	0	0.01	0.01	0.00	0	0.71	0.01	0.00	0	0	0.00	0	0	0.00	0.02	0	0.00	0.00	0.04	0	0	0.01	0	0.00	0	0	0.05	0	0.00	0	0	0.00	0	0	1
MN	0	0.00	0	0.02	0.05	0.03	0.00	0	0.00	0.02	0.01	0	0.03	0	0.05	0.02	0.00	0	0	0	0.01	0	0.02	0.46	0.01	0	0	0.01	0.00	0.02	0.00	0.00	0	0	0	0.08	0	0.00	0.02	0	0	0	0.00	0.02	0	0.01	0	0.02	0.07	0	0	1
MO	0	0	0	0	0.02	0.03	0.01	0	0	0.02	0.01	0	0.06	0	0.11	0.01	0.06	0	0	0.00	0	0	0.01	0.00	0.47	0	0	0	0	0.00	0	0	0	0	0.01	0.06	0	0	0.01	0	0.01	0	0.01	0.08	0	0	0	0	0	0	0	1
MS	0	0.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0	0	0	0	0	0.10	0	0.14	0	0	0	0.07	0	0	0	0	0.19	0	0	0	1
MT	0	0	0	0	0	0	0	0	0	0.12	0	0	0	0.09	0	0	0	0	0	0	0	0	0	0	0	0	0.65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.07	0	0	0.07	0	0	0	1
NC	0	0	0	0	0.03	0	0	0	0	0.08	0.09	0	0.01	0	0	0	0	0	0	0	0	0	0.05	0	0	0	0	0.56	0	0	0	0.02	0	0	0.02	0.01	0	0	0.00	0	0.08	0	0.01	0.01	0	0.04	0	0	0	0	0	1
ND	0	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0	0	0	0	0	0.57	0	0	0	0	0.24	0	0	0	0	0	0	0	0	0	0	0	0	0.03	0	0	0	0	0	0	0.04	0	0	1
NE	0	0	0	0	0.11	0.03	0	0	0	0.07	0.01	0	0.04	0	0.04	0.01	0.03	0	0	0	0.01	0	0.02	0.03	0.01	0	0	0	0	0.43	0	0	0	0	0.03	0.01	0	0	0.01	0	0	0	0	0.07	0	0.01	0.01	0	0.04	0	0	1
NH	0	0	0	0	0	0.11	0.18	0	0	0.18	0	0	0	0	0	0	0	0	0	0.23	0	0	0	0	0	0	0	0	0	0	0.18	0	0	0	0	0	0	0	0	0	0	0	0	0.11	0	0	0	0	0	0	0	1
NJ	0	0	0	0	0.09	0.02	0.02	0	0	0.05	0.00	0	0	0	0	0	0	0	0	0.02	0.01	0	0.01	0	0.00	0	0	0.02	0	0	0	0.42	0.01	0	0.08	0.01	0	0.01	0.13	0	0.01	0	0.01	0.06	0	0.01	0	0.01	0	0	0	1
NV	0	0	0	0	0.25	0.03	0	0	0	0.03	0	0.02	0	0	0.03	0.01	0	0	0	0.01	0	0	0	0	0	0	0	0.03	0	0	0	0	0.03	0.12	0.05	0.01	0.03	0	0.06	0	0	0	0	0.01	0	0.02	0	0.26	0	0	0	1
NY	0	0	0	0	0.06	0	0.05	0	0	0.02	0.03	0.01	0	0.00	0.01	0	0	0.01	0.01	0.02	0.00	0	0.01	0	0	0	0	0.02	0	0	0	0.06	0	0	0.55	0.02	0	0	0.06	0	0	0	0.02	0.02	0	0.01	0.01	0.01	0	0	0.01	1
OH	0	0	0.00	0.00	0.07	0.01	0	0	0	0.05	0.01	0	0	0	0.03	0.03	0.01	0.03	0	0.00	0.01	0	0.06	0	0.01	0.01	0	0.01	0	0	0	0.01	0	0	0.00	0.54	0	0	0.03	0	0	0	0.03	0.03	0.01	0.00	0	0	0	0.00	0	1
OK	0	0	0	0	0.08	0	0	0	0	0.17	0.01	0	0	0	0.02	0	0.01	0	0	0	0	0	0	0	0	0	0	0.01	0	0.03	0	0	0	0	0	0.01	0.16	0.01	0.01	0	0	0	0.02	0.43	0	0	0	0.04	0	0	0	1
OR	0	0	0	0.01	0.06	0.00	0	0	0	0.01	0	0.00	0	0.02	0	0.00	0	0	0	0	0	0	0	0	0	0	0	0.01	0.00	0	0	0	0	0.01	0	0	0	0.62	0	0	0	0	0.00	0.02	0	0.00	0	0.22	0	0	0	1
PA	0	0.01	0	0.01	0.03	0.01	0.01	0	0.03	0.05	0.05	0	0.00	0	0.01	0.00	0	0.01	0	0.01	0.01	0	0	0	0	0	0	0.02	0	0	0	0.07	0	0	0.04	0.01	0	0	0.47	0	0.03	0	0.00	0.05	0	0.01	0	0.02	0.00	0.03	0	1
RI	0	0	0	0	0	0	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0.09	0.02	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.88	0	0	0	0	0	0	0	0	0	0	0	1
SC	0	0	0	0.05	0.03	0	0	0	0	0.08	0.05	0	0	0	0.04	0	0	0.02	0.17	0	0	0	0	0	0.02	0.02	0	0.02	0	0	0	0.06	0	0	0	0.03	0	0	0	0	0.38	0	0.03	0.01	0	0	0	0	0	0	0	1
SD	0	0	0	0	0	0.02	0	0	0	0.16	0	0	0.07	0	0	0	0	0	0	0	0	0	0	0.06	0	0	0.01	0	0.08	0	0	0	0	0	0	0	0	0	0	0	0	0.46	0	0.04	0	0.06	0	0	0.02	0	0.01	1
TN	0	0.02	0.03	0	0.13	0	0	0	0	0.11	0.06	0	0	0	0.02	0	0	0.01	0	0	0	0.01	0	0.00	0	0.01	0	0.10	0	0.01	0	0.01	0	0	0.01	0.04	0	0	0	0	0.00	0	0.38	0.02	0.00	0.03	0	0	0	0	0	1
TX	0	0.00	0.00	0.01	0.13	0.01	0	0	0	0.06	0.02	0.00	0	0	0.02	0	0	0.01	0.02	0.00	0.00	0	0.02	0.00	0.01	0	0	0.02	0	0.00	0	0.01	0.01	0.01	0.02	0.01	0.01	0	0.01	0	0.00	0	0.00	0.58	0	0.00	0	0.01	0	0	0	1
UT	0	0	0	0.02	0.28	0.03	0	0	0	0.03	0	0.00	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0	0.01	0.04	0.01	0	0.03	0.03	0	0	0.01	0	0.10	0.33	0.03	0	0.04	0.00	0	0	1
VA	0	0	0	0.00	0.08	0.05	0	0.01	0	0.05	0.03	0	0	0	0.01	0	0	0.01	0	0.02	0.08	0	0.00	0.01	0	0	0	0.05	0	0	0	0.01	0	0	0.02	0.02	0	0	0.11	0	0	0	0.00	0.05	0	0.38	0	0.01	0	0.01	0	1
VT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0	0	0	0	0	0	0.16	0	0	0	0.04	0	0	0	0	0	0	0	0	0	0	0	0.74	0	0	0	0	1
WA	0	0	0	0.00	0.13	0.05	0	0	0	0.03	0	0	0	0	0	0	0	0	0	0.01	0.01	0	0	0.01	0	0.01	0	0	0	0	0	0.01	0	0.00	0.02	0.00	0	0.12	0	0	0	0	0	0.09	0	0	0	0.49	0.01	0	0	1
WI	0	0	0	0.01	0.02	0.01	0	0	0	0.04	0	0	0.02	0	0.11	0.00	0	0	0	0	0	0	0.08	0.14	0.03	0	0	0.01	0.01	0	0	0.00	0	0	0.00	0.02	0	0	0	0	0	0	0	0.01	0	0	0	0	0.48	0	0	1
WV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.13	0	0	0	0	0	0	0	0	0	0.08	0	0	0	0	0	0	0	0	0	0	0.15	0	0	0.27	0	0	0	0	0	0	0.05	0	0	0	0.32	0	1

For a chain whose home state is AL, for example, its aggregate economic variable is the weighted sum of the variable in the 50 states (plus DC) where the weights are given by the first row.
Weights larger than or equal to 0.10 are highlighted.

D Additional Descriptive Analyses

In this section, we repeat the regressions in Table 2 using an instrumental variable approach to address potential endogeneity problems with the key variable of interest: collateralizable housing wealth. We need to instrument for both collateralizable housing wealth and its product with employees. We construct our instrumental variables based on the elasticity estimates in Saiz (2010), which we aggregate from the MSA level to the state level to match the level of our data, using MSA population as weights in those cases where there is more than one MSA in a state as well as cases where an MSA is in more than one state. We then interact this instrumental variable with our measure of number of employees to obtain an instrument for (collateralizable housing wealth)×(employees). As in the paper, we link the macroeconomic data with the chain-level data using Matrix A. In these IV regressions, we can no longer include state fixed effects because the Saiz elasticity measure varies only cross-sectionally at the state level.

The results are similar: the estimated coefficient of collateralizable housing wealth is positive. However, the standard errors are larger, leading to statistically insignificant results when we instrument this variable. This is a usual consequence of using an instrument whose variability is limited – the Saiz instrument is strictly cross-sectional while the variable of interest varies both cross-sectionally and over time.[11] The estimated coefficients of capital needed and credit supply remain statistically significant.

E Details on the Log-likelihood Function

In this section, we derive the log-likelihood function (17). It consists of three components: the likelihood that chain i starts franchising in year F_i, $p_i(F_i|\boldsymbol{u}_i)$; the likelihood that this chain is in the sample $p_i(F_i \leq 2006|\boldsymbol{u}_i)$; and the likelihood of observing its two growth paths, for the number of company-owned and the number of franchised outlets, $p_i(n_{cit}, n_{fit}; t = F_i, ..., 2006|F_i; \boldsymbol{u}_i)$.

First, the likelihood of observing F_i conditional on chain i's unobservable component of the arrival rate and its unobservable profitability of opening a franchised outlet is

$$p_i(F_i|\boldsymbol{u}_i) = \sum_{t'=B_i}^{F_i} \left[\prod_{t=B_i}^{t'-1} (1-q_t) \cdot q_{t'} \cdot \prod_{t=t'}^{F_i-1} (1 - g(\boldsymbol{x}_{it}; \boldsymbol{u}_i)) \cdot g(\boldsymbol{x}_{iF_i}; \boldsymbol{u}_i) \right] \tag{E.1}$$

where t' represents when the chain starts to draw the entry cost from the non-degenerated log-normal distribution, q_t is the probability of this event in a specific year, and $g(\cdot,\cdot)$ is the conditional

[11] Moreover, the Saiz instrument is derived from data about metropolitan statistical areas whereas our measure of collateralizable wealth is for a chain's expected expansion area, which is typically even larger than a state. Even though the instrument is aggregated to the state level using population weights, and then to the greater area we are concerned about using the matrix mentioned previously, the resulting variable need not represent a good measure of elasticity at the more aggregate level.

Table D.1: Table 2 using IV

collateralizable housing wealth	1.332	1.249
	(1.397)	(1.400)
interest rate	-0.331	-0.359
	(0.570)	(0.557)
credit supply	0.526**	0.505**
	(0.231)	(0.234)
capital needed	-0.733**	-0.732**
	(0.353)	(0.351)
population	-0.091	-0.088
	(0.149)	(0.156)
employees	0.202	0.200
	(0.250)	(0.242)
(coll. housing wealth)×(employees)	-0.023	-0.034
	(0.060)	(0.055)
sector fixed effects		yes
R2	0.04	0.05
Observations	304	304

probability of entry into franchising, given in equation (12). As explained in the main body of the paper, $q_t = q_0$ in the year when the chain starts its business, denoted by B_i, (i.e., when $t = B_i$), and $q_t = q_1$ when $t > B_i$. Thus, the first summand in (E.1) (when $t' = B_i$) is $q_0 \cdot \prod_{t=B_i}^{F_i-1} (1 - g(\boldsymbol{x}_{it}; \boldsymbol{u}_i)) \cdot g(\boldsymbol{x}_{iF_i}; \boldsymbol{u}_i)$, the probability that chain i starts to draw the entry cost from the non-degenerated log-normal distribution from the very beginning, but chooses not to start franchising until year F_i. Similarly, the second summand in (E.1) (when $t' = B_i + 1$) is $(1 - q_0)q_1 \cdot \prod_{t=B_i+1}^{F_i-1} (1 - g(\boldsymbol{x}_{it}; \boldsymbol{u}_i)) \cdot g(\boldsymbol{x}_{iF_i}; \boldsymbol{u}_i)$, which is the probability that chain i starts to draw the entry cost from the non-degenerated log-normal distribution one year after it starts its business, but does not start franchising until year F_i. The sum of all such terms gives us the probability of starting franchising in year F_i.

Second, the likelihood of observing chain i in the sample, which requires that $F_i \leq 2006$, is the sum of the probability that chain i starts franchising right away ($F_i = B_i$), the probability that it starts one year later ($F_i = B_i + 1$), ..., and the probability that it starts in 2006, i.e.,

$$p_i (F_i \leq 2006|\boldsymbol{u}_i) = \sum_{F=B_i}^{2006} p_i (F|\boldsymbol{u}_i) . \tag{E.2}$$

Third, to derive the likelihood of observing the two growth paths ($n_{cit}, n_{fit}; t = F_i, ..., 2006$) of chain i conditional on its timing to enter into franchising, note that the number of company-owned outlets in year t is given by equation (15), copied below:

$$n_{cit} = n_{cit-1} - \text{exits}_{cit-1} + (\text{new outlets})_{cit}\,, \tag{E.3}$$

where $(n_{cit-1} - \text{exits}_{cit-1})$ follows a binomial distribution parameterized by n_{cit-1} and $1-\gamma$, the outlet exit rate; and $(\text{new outlets})_{cit}$ follows a Poisson distribution with mean $m_i p_{ac}(\boldsymbol{x}_{it}, \boldsymbol{u}_i)$ or $m_i p_{bc}(\boldsymbol{x}_{it})$ depending on whether the chain starts franchising before year t or not. Given that the mixture of a Poisson distribution and a binomial distribution is a Poisson distribution[12] and the sum of two independent Poisson random variables follows a Poisson distribution, n_{cit} follows a Poisson distribution with mean $\sum_{k=B_i}^{t} m_i p_c(\boldsymbol{x}_{ik}, \boldsymbol{u}_i)(1-\gamma)^{t-k}$, where $p_c(\boldsymbol{x}_{ik}, \boldsymbol{u}_i) = p_{bc}(\boldsymbol{x}_{ik})$ for $k < F_i$ and $p_c(\boldsymbol{x}_{ik}, \boldsymbol{u}_i) = p_{ac}(\boldsymbol{x}_{ik}, \boldsymbol{u}_i)$ for $k \geq F_i$. The likelihood of observing n_{cit} in the year the chain starts franchising (i.e., in the first year that we could observe this chain in the data) conditional on it starting franchising in year F_i is therefore

$$p_{n_{cit}|F_i}(\boldsymbol{u}_i) = \Pr\left(n_{cit}; \sum_{k=B_i}^{t} m_i p_c(\boldsymbol{x}_{ik}, \boldsymbol{u}_i)(1-\gamma)^{t-k}\right) \text{ for } t = F_i, \tag{E.4}$$

where $\Pr(\cdot; M)$ denotes the Poisson distribution function with mean M.

For subsequent years ($t = F_i + 1, ..., 2006$), we need to compute the likelihood of observing n_{cit} conditional on F_i as well as n_{cit-1}. According to equation (E.3), this conditional probability is the convolution of a binomial distribution ("$n_{cit-1} - \text{exits}_{cit-1}$" follows a binomial distribution with parameters n_{cit-1} and $1-\gamma$) and a Poisson distribution ("new outlets$_{cit}$" follows a Poisson distribution with mean $m_i p_{ac}(\boldsymbol{x}_{it}, \boldsymbol{u}_i)$):

$$\begin{aligned} & p_{n_{cit}|n_{cit-1},F_i}(\boldsymbol{u}_i) \\ &= \sum_{K=0}^{n_{cit-1}} \Pr(K|n_{cit-1}; 1-\gamma) \Pr\left((\text{new outlets})_{cit} = n_{cit} - K; m_i p_{ac}(\boldsymbol{x}_{it}, \boldsymbol{u}_i)\right), \end{aligned} \tag{E.5}$$

where K represents the number of outlets (out of the n_{cit-1} outlets) that do not exit in year $t-1$.

The conditional probabilities $p_{n_{fit}|F_i}(\boldsymbol{u}_i)$ and $p_{n_{fit}|n_{fit-1},F_i}(\boldsymbol{u}_i)$ can be computed analogously. Since the Poisson arrival events are independent events, the likelihood of observing chain i's growth

[12] If X follows a binomial distribution with parameters (M, p) and M itself follows a Poisson distribution with mean $\bar{M}$, then X follows a Poisson distribution with mean $\bar{M}p$.

path $p_i\left(n_{cit}, n_{fit}; t = F_i, ..., 2006 | F_i; \boldsymbol{u}_i\right)$ is the product of

$$
\begin{aligned}
& p_{n_{cit}|F_i}\left(\boldsymbol{u}_i\right), \text{ for } t = F_i; \\
& p_{n_{cit}|n_{cit-1},F_i}\left(\boldsymbol{u}_i\right), \text{ for } t = F_i + 1, ..., 2006; \\
& p_{n_{fit}|F_i}\left(\boldsymbol{u}_i\right), \text{ for } t = F_i; \\
& p_{n_{fit}|n_{fit-1},F_i}\left(\boldsymbol{u}_i\right), \text{ for } t = F_i + 1, ..., 2006.
\end{aligned}
\tag{E.6}
$$

In our likelihood function, we also handle missing data. For example, data in 1999 and 2002 were not collected.[13] When n_{cit-1} is not observable but n_{cit-2} is, we need to compute $p_{n_{cit}|n_{cit-2},F_i}$. Note that $n_{cit} = n_{cit-2} - \text{exits}_{cit-2} + (\text{new outlets})_{cit-1} - \text{exits}_{cit-1} + (\text{new outlets})_{cit}$, which can be rewritten as

$$
\begin{aligned}
& \text{outlets in } n_{cit-2} \text{ that do not exit before } t \\
& + \text{new outlets in } t-1 \text{ that do not exit before } t \\
& + \text{new outlets in } t,
\end{aligned}
$$

where "outlets in n_{cit-2} that do not exit before t" follows a binomial distribution with parameters $\left(n_{cit-2}, (1-\gamma)^2\right)$, "new outlets in $t-1$ that do not exit before t" follows a Poisson distribution with mean $m_i p_{ac}\left(\boldsymbol{x}_{it-1}, \boldsymbol{u}_i\right)(1-\gamma)$ and "new outlets in t" follows a Poisson distribution with mean $m_i p_{ac}\left(\boldsymbol{x}_{it}, \boldsymbol{u}_i\right)$. Therefore,

$$
\begin{aligned}
& p_{n_{cit}|n_{cit-2},F_i}\left(\boldsymbol{u}_i\right) \\
= & \sum_{K=0}^{n_{cit-2}} \Pr\left(K | n_{cit-2}; (1-\gamma)^2\right) \Pr\left((\text{new outlets})_{cit} = \right. \\
& \left. n_{cit} - K; m_i p_{ac}\left(\boldsymbol{x}_{it-1}, \boldsymbol{u}_i\right)(1-\gamma) + m_i p_{ac}\left(\boldsymbol{x}_{it}, \boldsymbol{u}_i\right)\right).
\end{aligned}
\tag{E.7}
$$

When more than one year of data is missing, we compute the corresponding conditional probability analogously. We then replace $p_{n_{cit}|n_{cit-1},F_i}$ and $p_{n_{fit}|n_{fit-1},F_i}$ in (E.6) by $p_{n_{cit}|n_{cit-2},F_i}$ and $p_{n_{fit}|n_{fit-2},F_i}$ when the observation of a year is missing, by $p_{n_{cit}|n_{cit-3},F_i}$ and $p_{n_{fit}|n_{fit-3},F_i}$ when data for two years are missing, etc.

[13] In addition to the two completely missing years, there are another 280 missing observations. Our results are robust to excluding the 89 firms with missing data for more than two years.

F Simulated Distributions of the Number of Outlets when Selection is Ignored

In our empirical model in Section 4, we endogenized the timing of entry into franchising. In this section, we show what would be the fit of the model, i.e., distributions of the predicted number of company-owned and franchised outlets, when the decision on the timing of entry into franchising is ignored. Specifically, in these simulations we take the observed waiting time in the data as exogenously given. The simulated distributions are shown in the right panels of Figure F.1. We also include the panels from Figures 2(b), 2(c) and 2(d), which show the distribution in the data and the simulated distribution taking selection into account, respectively, as the left and the middle panels in Figures F.1(a), F.1(b) and F.1(c), for comparison.

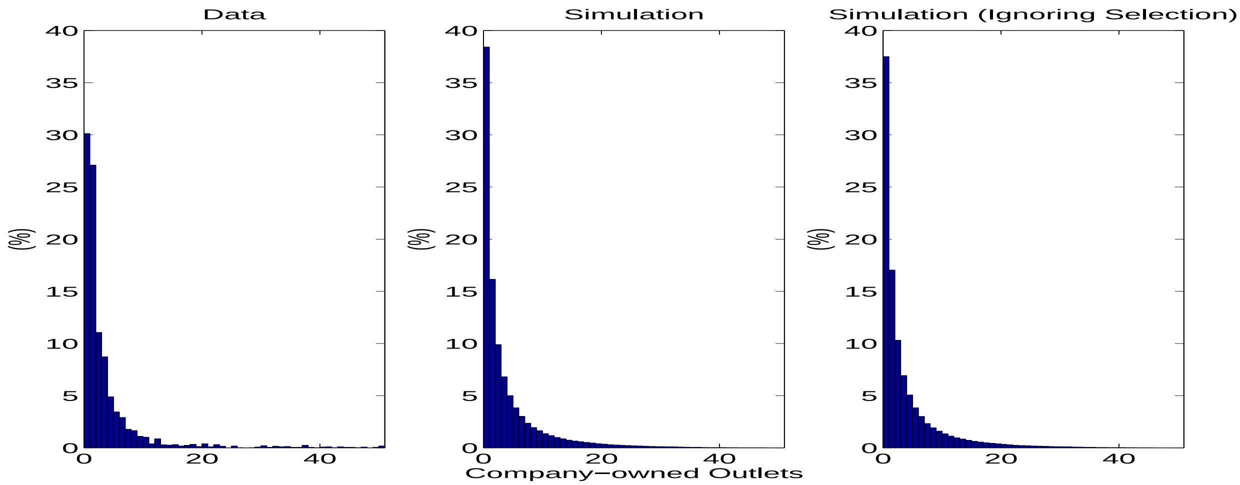
When we compare the middle panel of Figure F.1(a) (the simulated distribution of the number of company-owned outlets when selection is considered) and the right panel of the same figure (the simulated distribution when the timing of entry is ignored), we can see that these two distributions are very similar. This is because two effects are at play, and they apparently cancel each other out. On the one hand, chains that enter into franchising quickly tend to grow faster overall either because they are presented with more opportunities to open outlets or because outlets of these chains are more likely to be profitable. On the other hand, chains that enter fast into franchising are chains for which a franchised outlet is likely to be particularly profitable relative to a company-owned outlet. The latter effect shifts the distribution of the number of company-owned outlets to the left, while the first effect shifts the same distribution to the right.

Ignoring the timing of entry, however, makes a difference for the model fit in the number of franchised outlets and the fraction of franchised outlets. We can see this from the comparison of the middle panel and the right panel of Figure F.1(b) for the number of franchised outlets, and the same comparison in Figure F.1(c) for the fraction of franchised outlets. The simulated distribution of the number of franchised outlets when the entry decision is taken as exogenous (in the right panel) is shifted to the left from the simulated distribution where the entry decision is endogenized (in the middle panel). In particular, the percentage of observations with zero franchised outlets is over predicted by 16%. Similarly, the simulated proportion of firm/years with very high levels of franchised outlets is underestimated in Figure F.1(c).

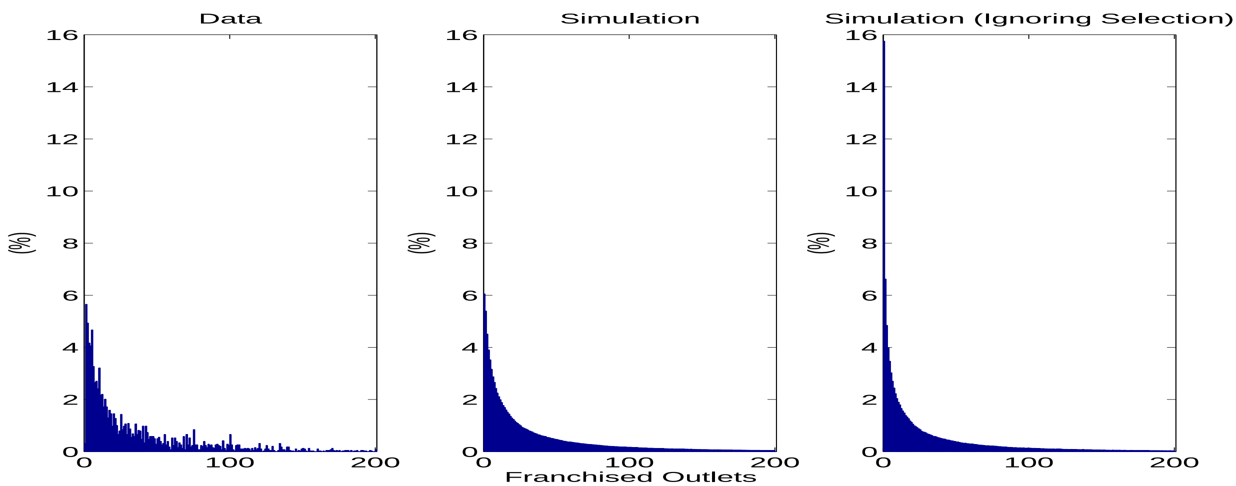
This is because when we simulate the distribution in the right panel, we draw the unobservable profitability of a franchised outlet from the unconditional distribution. So even if the draw is not in favor of a chain opening a franchised outlet when an opportunity arrives, the simulated number of franchised outlets corresponding to this draw, which is most likely to be very small, is included to compute the distribution. When the timing of entry is endogenized, however, a chain with unfavorable draws is likely to delay its entry into franchising, and therefore may not be included in the computation of the conditional distribution of the number of franchised outlets.

Figure F.1: Simulated Distributions of the Number of Outlets when Selection is Ignored

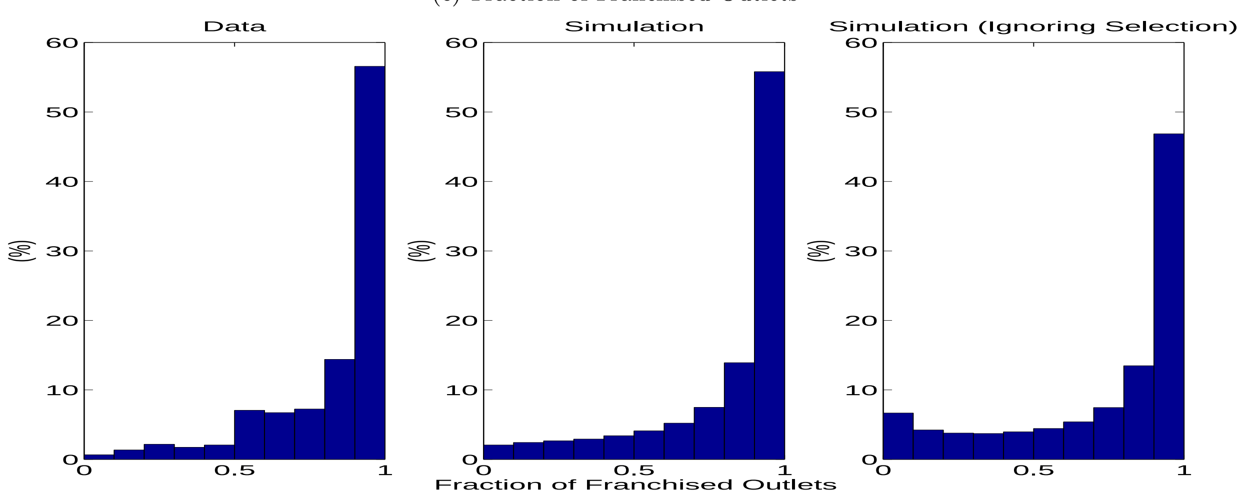
(a) Number of Company-owned Outlets



(b) Number of Franchised Outlets



(c) Fraction of Franchised Outlets



G Additional Robustness Analyses

In this section, we report results of two additional robustness analyses. In Section 5.4 in the paper, we show that our results are robust to including state fixed effects (additively). In the first additional robustness analysis (see Table G.1), we show that our results are also robust to including such state fixed effects in a non-additive fashion. Specifically, we estimate a model where we interact our key variable of interest (collateralizable housing wealth in the relative profitability of franchising) with state dummy variables. In other words, we allow the coefficient of collateralizable housing wealth to be state-specific. Note that in this robustness analysis, we do not include the "additive" state fixed effects so as to keep the number of parameters manageable.[14] In Table G.1, we report the estimated average effect of collateralizable housing wealth in the relative profitability of franchising (averaged across states) and its standard error. The estimated average effect is 0.224.[15] In comparison, the estimated effect of collateralizable housing wealth in the baseline results (where the coefficient is not state-specific) is 0.221. The table also shows that estimates of all other parameters are robust as well.

In the second additional robustness analysis, we re-estimate our empirical model with the National Foundation of Independent Business (NFIB)' measure of credit supply replaced by two alternative measures from the Federal Reserve's "Senior Loan Officer Opinion Survey on Bank Lending Practices" (SLOOS). These two measures are: (1) the net percentage of domestic banks loosening standards for C&I loans to small firms, and (2) the net percentage of domestic banks decreasing their collateral requirements for small firms. The estimation results are reported in Table G.2. While the estimates of other parameters are largely unchanged, the estimated effects of the SLOOS measures of loan availability are not statistically significant, probably for reasons explained in Footnote 18.

[14]In a separate estimation, we include regional fixed effects (Northeast, Midwest, South and West) both additively and interacted with collateralizable housing wealth. Our results remain robust.

[15]The standard deviation of the estimated state-specific coefficient (across states) is 0.116.

Table G.1: Robustness Analysis: Allowing the Coefficient of Collateralizable Housing Wealth in the Relative Profitability of a Franchised Outlet to be State Specific

	est.	s.e.
Log of opportunity arrival rate		
constant	3.047***	0.015
std. dev.	1.304***	0.024
General profitability		
constant	-3.561***	0.045
population	0.258***	0.006
per-capita state product	0.010***	0.001
collateralizable housing wealth	-0.049***	0.008
Relative profitability of a franchised outlet		
collateralizable housing wealth	0.224***	0.0004
employees	0.008*	0.005
(coll. housing wealth)×(employees)	0.018***	0.001
interest rate	-0.041***	0.003
capital needed	-0.514***	0.021
credit supply	1.326***	0.029
population	0.007***	0.002
business products & services	-0.040	0.073
restaurant	-0.057	0.069
home services	1.166***	0.079
go to services	0.699***	0.070
auto; repair	1.351***	0.072
constant (retailer)	2.239***	0.107
std. dev.	2.452***	0.029
Outlet exit rate	0.317***	0.001
Entry cost		
mean parameter	3.104***	0.275
std. dev. parameter	0.622***	0.211
probability q_0	0.175***	0.026
probability q_1	0.186***	0.016

* indicates 90% level of significance. *** indicates 99% level of significance.

Table G.2: Robustness Analysis: Using the SLOOS Measures for Loan Supply

	(1)		(2)	
	est.	s.e.	est.	s.e.
Log of opportunity arrival rate				
constant	3.306***	0.022	3.367***	0.022
std. dev.	1.197***	0.035	1.314***	0.037
General profitability				
constant	-3.581***	0.080	-3.540***	0.080
population	0.194***	0.008	0.200***	0.009
per-capita state product	0.008***	0.001	0.008***	0.002
collateralizable housing wealth	-0.031***	0.013	-0.051***	0.013
Relative profitability of a franchised outlet				
collateralizable housing wealth	0.227***	0.012	0.294***	0.015
employees	-0.003	0.007	-0.003	0.011
(coll. housing wealth)×(employees)	0.016***	0.001	0.005***	0.002
interest rate	-0.165***	0.005	-0.132***	0.006
capital needed	-0.470***	0.056	-0.284***	0.043
credit supply	-0.0004	0.004	-0.0002	0.006
population	0.004	0.004	0.003	0.003
business products & services	-0.011	0.136	-0.034	0.138
restaurant	-0.100	0.114	0.255**	0.120
home services	0.703***	0.099	0.507***	0.133
go to services	0.007	0.136	0.234**	0.127
auto; repair	1.533***	0.146	1.078***	0.266
constant (retailer)	2.382***	0.145	1.677***	0.160
std. dev.	2.573***	0.076	2.615***	0.080
Outlet exit rate	0.342***	0.002	0.344***	0.002
Entry cost				
mean parameter	2.694***	0.136	2.839***	0.333
std. dev. parameter	0.424***	0.017	0.528*	0.381
probability q_0	0.168***	0.022	0.173***	0.032
probability q_1	0.232***	0.024	0.230***	0.027

* indicates 90% level of significance. ** indicates 95% level of significance. *** indicates 99% level of significance.