Online Appendix to "Indirect Effects of Access to Finance"*

Jing Cai
University of Maryland, NBER and BREAD

Adam Szeidl
Central European University and CEPR

January 12, 2024

A Additional theoretical and empirical results

A.1 Theory: Assumption, proofs, calculations, extensions

A.1.1 Assumptions about firm dynamics

We say that a distribution over a finite-dimensional Euclidean space is smooth if it has a continuous density. Let k_m be a vector of market-specific characteristics which summarizes the initial distribution of firm quality and productivity in the market, and is drawn from a smooth distribution.

Firm level log quality and productivity absent the intervention evolve as $\tilde{x}_i^t = b_i^t + g^t + \eta_{xi}^t$. Here b_i^t is a two-dimensional vector of firm level latent characteristics, g^t is 2-dimensional vector representing a deterministic time trend common across all firms, and η_{xi}^t is a 2-dimensional vector i.i.d. across firms and over time and drawn from a smooth distribution. The initial value of firm-level latent characteristics, b_i^0, is drawn for each firm i from a smooth distribution $B^0(k_m)$ that depends continuously on k_m. Latent characteristics evolve according to a Markov process, $b_i^t = B(b_i^{t-1}, \eta_{bi}^t, \eta_{bm}^t)$ where the η_{bi}^t vector of firm-level shocks are i.i.d. across firms and over time and drawn from a smooth distribution, while the η_{bm}^t vector of market-level shocks are i.i.d. across markets and over time and drawn from a smooth distribution. The B function is differentiable with a continuous derivative.

We assume that all these variables have well-defined means and variances. We also assume that the continuum law of large numbers holds (Sun 2006), so that cross-sectional averages in a

* Emails: cai516@umd.edu, szeidla@ceu.edu.

realization are equal to the analogously-defined expected values of the same variables. We use the imprecise but convenient notation that the expectations operator denotes both the expectation in the probabilistic model and the average across firms and over time in a realization. The former is useful for some of the derivations, and the latter is useful for analyzing the identifying assumptions.

A.1.2 Proof of Proposition 1

We first solve the model without making any assumptions about the treatment. Normalize the price of H to one. Start with the optimal allocation in market m given budget E_m. Maximizing (1) subject to the budget constraint $\int_{i\in m} P_i Q_i di = E_m$ implies

$$Q_i = Q_m h_m^{\sigma-1} P_i^{-\sigma} l_m^{-\sigma}$$

for $i \in m$ where l_m is the multiplier on the constraint. Expressing $(h_i Q_i)^{1-1/\sigma}$ from this and integrating over i gives

$$\frac{1}{l_m} = \left(\int_{i\in m} \left(\frac{P_i}{h_i} \right)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}}$$

which is the quality-adjusted price index for market m, and which we denote by P_m. It then follows directly that

$$\frac{h_i Q_i}{Q_m} = \left(\frac{P_i/h_i}{P_m} \right)^{-\sigma} \tag{A1}$$

for $i \in m$ so that the quality-adjusted relative quantity relates with elasticity $-\sigma$ to the quality-adjusted relative price of a product in market m. Moreover rearranging this implies that $\int_{i\in m} P_i Q_i di = P_m Q_m$, justifying our definition of P_m.

Now consider maximization across markets. Given a budget E that we assume is sufficiently large that in the unconstrained optimum there is positive consumption of the numeraire, the consumer's objective can be rewritten as

$$E + \int Q_m^{1-1/\theta} dm - \int P_m Q_m dm.$$

Maximizing this yields the first-order condition

$$(1 - 1/\theta)Q_m^{\frac{-1}{\theta}} = P_m.$$

From this we can express the demand for the composite good m as

$$Q_m = \left(\frac{\theta}{\theta - 1}\right)^{-\theta} P_m^{-\theta}. \tag{A2}$$

From these and the previous equations, given prices P_i and qualities h_i, all quantities can be expressed.

Now consider firms' price-setting. Because a firm is small relative to the market, it does not take into account its impact on the price indices. As a result, maximizing profits $P_i Q_i - w L_i$ subject to the demand curve (A1) yields the usual constant markup

$$P_i = \frac{\sigma}{\sigma - 1} \cdot \frac{w}{\omega_i}.$$

From (A1) and (A2), firm revenue can be written as

$$P_i Q_i = \left(\frac{\theta}{\theta - 1}\right) \left(\frac{P_i}{h_i}\right)^{1-\sigma} P_m^{\sigma-\theta} \tag{A3}$$

for $i \in m$. This equation can be used to characterize revenue both with and without the treatment.

Effect of loan program on revenue. As (A3) shows, the loan program can affect firm revenue in two ways: through the firm's quality-adjusted price P_i/h_i and through the market's quality-adjusted price index P_m. Given our assumption that the loan acts by changing quality and cost, so that for a borrower firm i, $\Delta \log h_i = \gamma_h$ and $\Delta \log \omega_i = \gamma_\omega$, the program's effect on the firm's log quality-adjusted price is $\Delta \log(P_i/h_i) = -(\gamma_\omega + \gamma_h) = -\gamma$.

To characterize the program's effect on the price index P_m, note that

$$P_m = \left(\int_{i \in m} \left(\frac{P_i}{h_i} \right)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}} = \left(\int_{i \in m} \left(\frac{\tilde{P}_i}{\tilde{h}_i} \right)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}} \cdot (1 - \lambda_m Z_m + \lambda_m Z_m e^{\gamma(\sigma-1)})^{\frac{1}{1-\sigma}} \quad \text{(A4)}$$
$$= \tilde{P}_m \cdot (1 - \lambda_m Z_m + \lambda_m Z_m e^{\gamma(\sigma-1)})^{\frac{1}{1-\sigma}}$$

where we used that borrowers represent a share Z_m of firms in market m, that $(\tilde{P}_i^t / \tilde{h}_i^t)^{1-\sigma}$ is proportional to revenue and—absent the treatment—treated firms' revenue is on average λ_m times that of the average firm, and that the loan multiplies $\tilde{P}_i / \tilde{h}_i$ by $e^{-\gamma}$. It follows that

$$\Delta \log P_m = \frac{\log(1 - \lambda_m Z_m + \lambda_m Z_m e^{\gamma(\sigma-1)})}{1 - \sigma}. \quad \text{(A5)}$$

At $\gamma = 0$ the term on the right-hand side equals zero and has derivative with respect to γ of $-\lambda_m Z_m$. Thus to a first order approximation

$$\Delta \log P_m \approx -\gamma \lambda_m Z_m.$$

Substituting this into (A3) implies that the change in firm revenue in response to the treatment is

$$\Delta P_i Q_i \approx (\sigma - 1)\gamma \cdot B_i - (\sigma - \theta)\gamma \lambda_m \cdot Z_m \quad \text{(A6)}$$

for $i \in m$, as claimed.

A.1.3 Proof of Proposition 2

Take-up. We use the notation that for any variable X, the conditional expectation $E[X|t, m]$ means conditioning on characteristics specific to m (that is, the realized k_m) and t (that is, market-level shock realizations up to and including t) that do not include the treatment assignment. In contrast, the conditional expectation $E[X|t, m, T_i, S_m]$ means also conditioning on the treatment assignment (T_i, S_m). Let $\bar{x}_i = (\tilde{x}_i^1, \tilde{x}_i^s)$ denote the vector of firm characteristics absent the treatment that

influence whether firm i borrows either in period 1 due to the treatment or in period s due to diffusion. The probability that a firm with characteristics $\bar{x}_i$ in market m is a borrower in period t, conditional on its treatment status, is

$$E[B_i^t|t, m, \bar{x}_i, T_i, S_m] = T_i \cdot F^T(\tilde{x}_i^1) \cdot 1_{\{t \geq 1\}} + \phi(1 - T_i)S_m \cdot F^D(\tilde{x}_i^s) \cdot 1_{\{t \geq s\}}. \tag{A7}$$

Taking the conditional expectation with respect to T_i and S_m in period t implies that for $t \geq s$,

$$E[B_i^t|t, T_i, S_m] = \mu^T \cdot T_i + \phi\mu^D \cdot (1 - T_i)S_m.$$

This follows because the expectation that a firm borrows if treated, $F^T(\tilde{x}_i^1)$, is independent of its realized treatment assignment (T_i, S_m) and hence equal to μ^T; and the expectation that a firm borrows if reached by diffusion, $F^D(\tilde{x}_i^s)$, is independent of its realized treatment assignment (T_i, S_m) and hence equal to μ^D. It follows that

$$B_i^t = \mu^T \cdot T_i + \phi\mu^D \cdot (1 - T_i)S_m + \eta_i^t$$

where $E[\eta_i^t|T_i, S_m] = 0$, implying the orthogonality condition claimed in the Proposition.

Averaging equation (A7) across all firms in market m implies

$$E[Z_m^t|t, m, S_m] = S_m E[F^D(x_i^1)|m] \cdot 1_{\{t \geq 1\}} + \phi(1 - S_m)S_m E[F^T(x_i^s)|m] \cdot 1_{\{t \geq s\}},$$

and because $Z_t^m = E[Z_m^t|t, m, S_m]$, we can write

$$Z_m^t = S_m \cdot \mu_m^T \cdot 1_{\{t \geq 1\}} + \phi(1 - S_m)S_m \cdot \mu_m^D \cdot 1_{\{t \geq s\}} \tag{A8}$$

where $\mu_m^T = E[F^D(x_i^1)|m]$ and $\mu_m^D = E[F^T(x_i^s)|m]$ are the average take-up rates due to treatment and to diffusion in market m.

Firm revenue. We can write

$$\log P_i^t Q_i^t = \log \tilde{P}_i^t \tilde{Q}_i^t + \Delta \log P_i^t Q_i^t \tag{A9}$$

where, as before, tilde denotes outcomes absent the intervention. Rewriting the second term by applying Proposition 1 to period t, firm revenue is

$$\log R_i^t \approx \log \tilde{R}_i^t + (\sigma - 1)\gamma \cdot B_i^t - (\sigma - \theta)\gamma \lambda_m^t \cdot Z_m^t.$$

As a first step towards establishing Part 2 of the Proposition, we show that $\log \tilde{R}_i^t$ is orthogonal to the residualized instruments, which are defined by projecting the instruments $Post^t \cdot T_i$ and $Post^t \cdot S_m$ on the firm fixed effects and $Post^t$, and taking the residual. It is straightforward to verify that these residualized instruments are $Post^t \cdot (T_i - \bar{S})$ and $Post^t \cdot (S_m - \bar{S})$ where $\bar{S} = E[S_m] = E[T_i]$ is the average treatment intensity. Indeed, these are obtained by transforming the instruments with a linear function of $Post^t$ and are orthogonal to the firm fixed effects and $Post^t$. The claimed orthogonality conditions are then

$$E[\log \tilde{R}_i^t \cdot Post^t \cdot (T_i - \bar{S})] = 0$$

and

$$E[\log \tilde{R}_i^t \cdot Post^t \cdot (S_m - \bar{S})] = 0$$

which both hold because T_i and S_m are randomly assigned independently of the $\tilde{R}_i^t$ realizations.

In our second step of the proof, we decompose $\log \tilde{R}_i^t$ into the sum of a component spanned by the firm effects and $Post^t$, and another component, denoted ν_i^t, orthogonal to the firm effects and $Post^t$. Because $\log \tilde{R}_i^t$ is orthogonal to the residualized instruments, and by construction so is the first component, it follows that so is ν_i^t. But then ν_i^t, which is orthogonal to the firm effects and $Post^t$ is also orthogonal to the instruments themselves. Thus we can write

$$\log R_i^t \approx f_i + \kappa \cdot Post^t + (\sigma - 1)\gamma \cdot B_i^t - (\sigma - \theta)\gamma \lambda_m^t \cdot Z_m^t + \nu_i^t \tag{A10}$$

where $E[\nu_i^t \times (1_j, Post^t, Post^t \cdot T_i, Post^t \cdot S_m] = 0$ holds for all j. This is almost our desired equation, except that it involves λ_m^t rather than λ.

Substituting average selection. For any fixed λ rewrite the equation as

$$\log P_i^t Q_i^t \approx f_i + \kappa \cdot Post^t + (\sigma - 1)\gamma \cdot B_i^t - (\sigma - \theta)\gamma\lambda \cdot Z_m^t + (\sigma - \theta)\gamma(\lambda - \lambda_m^t) \cdot Z_m^t + \nu_i^t.$$

We show below that with appropriate choice of λ the new term $(\lambda - \lambda_m^t) \cdot Z_m^t$ is orthogonal to the residualized instruments. This implies that $(\lambda - \lambda_m^t) \cdot Z_m^t$ can be decomposed into a component spanned by the firm and time effects and another component, denoted u_i^t, orthogonal to the firm and time effects and the instruments. It then follows that after replacing λ_m by λ, and changing the fixed effects, the orthogonality conditions for the IV will hold with the new error term $\varepsilon_i^t = u_i^t + \nu_i^t$.

Since the residualized instruments are $Post^t \cdot (T_i - \bar{S})$ and $Post^t \cdot (S_m - \bar{S})$, the required orthogonality conditions are

$$E[(\lambda - \lambda_m^t) Z_m^t \cdot (T_i - \bar{S}) Post^t] = 0$$

and

$$E[(\lambda - \lambda_m^t) Z_m^t \cdot (S_m - \bar{S}) Post^t] = 0.$$

Because in the first condition all terms except T_i depend only on m and t, and because the mean of T_i in market m is S_m, the second condition implies the first:

$$E[E[(\lambda - \lambda_m^t) Z_m^t \cdot (T_i - \bar{S}) Post^t] | t, m, S_m] = E[(\lambda - \lambda_m^t) Z_m^t \cdot (S_m - \bar{S}) Post^t].$$

The second condition would hold if

$$\lambda = \frac{E[\lambda_m^t \cdot Z_m^t (S_m - \bar{S}) Post^t]}{E[Z_m^t \cdot (S_m - \bar{S}) Post^t]}, \tag{A11}$$

which gives us a set of candidate weights

$$a_m^t = E[Z_m^t (S_m - \bar{S}) Post^t | t, m]$$

where defining the weights as conditional expectations does not change the weighted average because λ_m^t depends only on characteristics specific to t and m. This definition of λ ensures that the orthogonality conditions hold. We also need to verify that a_m^t are non-negative. To do this, recall from (A8) that

$$Z_m^t = S_m \mu_m^T \cdot 1_{\{t \geq 1\}} + \phi(1 - S_m) S_m \mu_m^D \cdot 1_{\{t \geq s\}}.$$

Using this, we can rewrite a_m^t as

$$a_m^t = \mu_m^T \cdot 1_{\{t \geq 1\}} E[S_m(S_m - \bar{S})] + \phi \mu_m^D \cdot 1_{\{t \geq s\}} E[(1 - S_m) S_m (S_m - \bar{S})] \tag{A12}$$

where we used that the distribution of S_m conditional on t and m is the same as its unconditional distribution due to the random assignment. Now note that $E[S_m(S_m - \bar{S})]$ is the variance of S_m and hence non-negative. Given this, our assumption that $\mu_m^T > \phi \mu_m^D$ implies that the sign of a_m^t when $t \geq 1$ is not lower than the sign of

$$E[(S_m + (1 - S_m) S_m) \cdot (S_m - \bar{S})],$$

which is the covariance between $S_m + (1 - S_m) S_m$ and S_m, and since both are non-decreasing functions of S_m, is non-negative.

A.1.4 Proof of Proposition 3

We first develop a characterization of welfare in any period that is valid with or without the treatment. Let $E_d = \int_m P_m Q_m dm$ denote expenditure on the differentiated products and Π denote total profits by all firms producing differentiated products. The consumer's budget constraint is

$$wL + \Pi = E = H + E_d. \tag{A13}$$

We assume throughout that L is sufficiently large so that $H > 0$. Then the consumer's maximized utility, or welfare, can be written as

$$wL + \Pi + \int Q_m^{1-1/\theta} dm - \int P_m Q_m dm.$$

Since L is fix, the treatment can affect welfare by changing the producer surplus Π or the consumer surplus

$$CS = \int Q_m^{1-1/\theta} dm - \int P_m Q_m dm.$$

From (A2) we can express the consumer surplus as

$$CS = \left(\frac{\theta}{\theta-1}\right)^{1-\theta} \int_m P_m^{1-\theta} dm - \left(\frac{\theta}{\theta-1}\right)^{-\theta} \int_m P_m^{1-\theta} dm = \frac{1}{\theta-1} \left(\frac{\theta}{\theta-1}\right)^{-\theta} \int_m P_m^{1-\theta} dm.$$

To use this formula for period t, recall that

$$P_m^t = \tilde{P}_m^t (1 - \lambda_m^t Z_m^t + \lambda_m^t Z_m^t e^{\gamma(\sigma-1)})^{\frac{1}{1-\sigma}} \tag{A14}$$

thus

$$\Delta CS^t = \frac{1}{\theta-1} \left(\frac{\theta}{\theta-1}\right)^{-\theta} \int_m (\tilde{P}_m^t)^{1-\theta} [(1 - \lambda_m^t Z_m^t + \lambda_m^t Z_m^t e^{\gamma(\sigma-1)})^{\frac{1-\theta}{1-\sigma}} - 1] dm$$

and differentiating the last term with respect to γ at $\gamma = 0$ gives $(\theta-1)\lambda_m^t Z_m^t$ and hence

$$\Delta CS^t \approx \frac{1}{\theta-1} \left(\frac{\theta}{\theta-1}\right)^{-\theta} \int_m (\tilde{P}_m^t)^{1-\theta} (\theta-1)\gamma \lambda_m^t Z_m^t dm = \gamma \int_m \tilde{R}_m^t \lambda Z_m^t dm.$$

Finally, using (A8), for $t \geq s$ we can write

$$\Delta CS^t \approx \int_m [S_m \mu_m^T + \phi(1 - S_m) S_m \mu_m^D] \lambda_m^t \gamma \tilde{R}_m^t dm.$$

A.1.5 Return on capital calculation

We compute the private and social return to capital, and decompose the social return, in two steps. First we use our estimate to measure yields earned at the midline and endline; and second, we

make assumptions about the dynamics of these yields and derive the implied rates of return.

We measure the average private yield in the midline and endline with the profit direct effect IV estimate ζ_Π as

$$\textit{Private yield} = \frac{\zeta_\Pi}{\textit{Avg loan}} + \textit{Interest rate} - \textit{Default rate}.$$

The first term is the yield accumulating to borrowers as profits, while the second and third terms measure the net yield accumulating to the bank. To compute the social yield, this private yield needs to be adjusted by business stealing and the consumer surplus. We measure the contribution of business stealing with the profit indirect effect IV estimate ξ_Π as

$$\textit{Business stealing yield} = \frac{\xi_\Pi}{\textit{Avg loan}}$$

where the business stealing effect of the loan on profits is normalized by the average loan size. And we measure the contribution of the consumer surplus as

$$\textit{Consumer surplus yield} = \frac{\frac{\zeta_R}{\sigma-1}\bar{R}_C}{\textit{Avg loan}}$$

where the gain in consumer surplus from the loan is normalized by loan size.

In our second step, we assume that the yield in year 1 is zero, the yield in year 2 is as computed above and denoted by y, and the yield in year $t \geq 2$ is $y(1-d)^{t-2}$ where d is the depreciation rate. We then compute the internal rate of return by noting that with a discount rate ρ, the present discounted value of this payment stream is

$$\frac{y}{1+\rho}\frac{1}{\rho+d}$$

which needs to be equal to 1 if ρ is the internal rate of return. This yields a quadratic equation that gives $\rho = [(4y+1+d^2-2d)^{1/2}-d-1]/2$. We solve for ρ when y is the private yield; when y is the private yield plus the business stealing yield; and when y is the private yield plus the business stealing yield plus the consumer surplus yield. The resulting ρ values, and the increments between them, give us the private return, the additional contribution of business stealing, the additional

contribution of consumer surplus, and the social return. We conduct this calculation using an annual depreciation rate of $d = 0.10$.

A.1.6 Exit

Our main theoretical analysis assumed that all firms stay alive and stay in the sample. But in practice some firms leave because they exit the market, and some firms do not answer the survey even if they stay in the market. We now discuss the impacts of exit and attrition on our empirical approach and welfare calculations.

Empirical design. We first derive conditions under which our empirical design remains valid in the presence of exit and attrition. Let A_i^t be an indicator for whether firm i is in the sample in period t. Thus $A_i^t = 0$ represents both exit and attrition. Let $\bar{\varepsilon}_i^t$ denote the residualized variable obtained by projecting ε_i^t on the firm effects and $Post$ conditional on $A_i^t = 1$, and taking the residual. Then the identifying condition for (5), conditional on observations with $A_i^t = 1$, becomes

$$E[\bar{\varepsilon}_i^t \times (Post^t \cdot T_i, Post^t \cdot S_m) | A_i^t = 1] = 0.$$

In words, the firm shocks, net of firm and time fixed effects, should be orthogonal to the right-hand side variables of our regression even in the subset of firms remaining in sample.

As a balance test for this condition, we can test whether firm characteristics at baseline are orthogonal to the right-hand side variables in the subset of firms remaining in sample:

$$E[\varepsilon_i^0 \times (T_i, S_m) | A_i^\tau = 1] = 0.$$

Table A3 confirms this for several key firm characteristics, both for firms that remain in sample until 2016 and until 2020. This result provides evidence against exit or attrition inducing bias in our estimates.

Welfare effects. Conditional on the coefficient estimates being correct, attrition does not affect the welfare calculations as long as the firms that do not answer the survey behave as described by our model. But exit is outside the model, and we need to incorporate it to characterize its impact

on welfare.

Motivated by the result that the treatment reduces the probability of exit, assume that each period $t \geq 1$ a share x_B of borrowers and a share x_N of non-borrowers exit. Assume that exiting borrowers and exiting non-borrowers are drawn randomly from the set of borrowers and non-borrowers in the market, respectively. This assumption is consistent with the data: at baseline the average revenue of exiting firms is not significantly different from that of all firms, and the average revenue of exiting borrowers is not significantly different from that of all borrowers.[1] For simplicity we assume that the exit rate of firms borrowing due to information diffusion—which only borrow in period $s \geq 1$—is already x_B starting from period 1.

Since we model exit, we also need to model entry. We make the simple assumption that exiting firms are replaced by firms which are drawn from the distribution of firms in the market absent the intervention. Since the exit rate is lower for borrowers ($x_B < x_S$), this assumption implies that having treated competitors reduces the entry rate somewhat. But it does not reduce the entry rate by too much, because the total number of firms in the market is held fixed. To keep the analysis conservative, we assume that replacement firms do not borrow. Finally, for ease of notation, in the derivation below we ignore depreciation, but it can be easily added by discounting γ with powers of $1 - d$ where d is the depreciation rate.

Now consider an intervention in which a share S of firms are treated in a set of markets M. We abstract away from cross-market heterogeneity. Similarly to the main text, assume the market-level selection coefficients λ_m^t are all identical to some value λ^0, but, differently from the main text, allow λ^0 to be different from one. Let $Z = S\mu^T + S(1-S)\phi\mu^D$ denote the share of borrowers in the market absent exit.

In a period $t \geq s$ we can write the market level price index as

$$P_m^t = \tilde{P}_m^t (1 + \lambda^0 (e^{\gamma(\sigma-1)} - 1) Z (1 - x_B)^t + [(1 - x_B)^t - (1 - x_N)^t] Z (\lambda^0 - 1))^{\frac{1}{1-\sigma}}.$$

There are two novelties relative to the analogous expression (A14) in the proof of Proposition

[1] When borrowers are larger than the average firm, our assumption would imply that exiting borrowers are larger than the average exiting firm, which seemingly contradicts with our finding that treated and untreated surviving firms are similar at baseline. But borrower firms are only slightly and insignificantly larger than the average firm.

3. First, the share of borrowers in period t is no longer Z, but is instead $Z(1-x_B)^t$ due to exit. These firms charge a quality-adjusted price which is reduced by factor e^{γ}. Second, a share $[(1-x_B)^t-(1-x_N)^t]Z$ of firms survived to period t because of borrowing. These firms are on average λ^0 times as large as the average firm, explaining the last term in the expression.

Using the previous formula, we take a first-order approximation of the impact of the intervention on the price index raised to $1-\theta$, around the point $\gamma=0$ and $\lambda^0=1$:

$$(P_m^t)^{1-\theta}-(\tilde{P}_m^t)^{1-\theta}\approx(\tilde{P}_m^t)^{1-\theta}\frac{1-\theta}{1-\sigma}\left[Z(1-x_B)^t(\sigma-1)\gamma+[(1-x_B)^t-(1-x_N)^t]Z(\lambda^0-1)\right].$$

This implies that

$$\begin{aligned}\Delta CS^t&=\frac{1}{\theta-1}\left(\frac{\theta}{\theta-1}\right)^{-\theta}\int_m(P_m^t)^{1-\theta}-(\tilde{P}_m^t)^{1-\theta}dm\\&\approx\left(\frac{\theta}{\theta-1}\right)^{-\theta}\int_m(\tilde{P}_m^t)^{1-\theta}dm\cdot\left[Z(1-x_B)^t\gamma+\frac{[(1-x_B)^t-(1-x_N)^t]Z(\lambda^0-1)}{\sigma-1}\right]\\&=\int_m\tilde{R}_m^t dm\cdot\left[Z(1-x_B)^t\gamma+\frac{[(1-x_B)^t-(1-x_N)^t]Z(\lambda^0-1)}{\sigma-1}\right].\end{aligned}$$

Finally, approximating revenue absent the intervention with revenue in control markets $\bar{R}_C$ yields

$$\frac{\Delta CS^t}{\sum_M n_m}\approx\frac{\zeta_R}{\sigma-1}Z\cdot\bar{R}_C(1-x_B)^t+[(1-x_B)^t-(1-x_N)^t]Z\frac{\lambda^0-1}{\sigma-1}\bar{R}_C. \tag{A15}$$

In this formula, the first term is that in (11) modified by $(1-x_B)^t$: the reduction in quality-adjusted prices is coming only from borrowers, whose share declines with exit. The second term is new and represents the impact of the reduction in the exit rate induced by borrowing. This reduction implies that a share $(1-x_B)^t-(1-x_N)^t$ of borrowers are saved from being replaced by average firms. To the extent that these firms have above-average revenue (λ^0-1), their quality adjusted price must be lower ($1/(\sigma-1)$ is the elasticity of price to revenue), so that their survival reduces the quality-adjusted price index.

Finally consider the producer surplus. A back-of-the envelope approach to incorporate the

impact of exit is to compute

$$\frac{\Delta PS}{\sum_M n_m} \approx (\zeta_\Pi + \xi_\Pi) Z (1 - x_B)^t. \tag{A16}$$

The novelty is $(1 - x_B)^t$: the share of borrowers—which generates both the direct and business stealing effect—is declining due to exit. This approach ignores the possible effect that the survival of larger-then-average borrowers may increase the producer surplus, and that the presence of those borrowers may intensify competition and reduce the producer surplus. Because borrowers are only slightly and insignificantly larger, we ignore these forces.

To operationalize these formulas, assume that $\lambda^0 = 1$, and because the share of borrowers that exit by endline is 1.95%, that $x_B = 0.07$ per year. Then the impact of exit on the producer surplus and the consumer surplus by the endline is less than 2 percent. We ignore this effect in the text.

Return to capital. To adjust the calculation of return to capital to exit, we need to bring back depreciation. Formally, depreciation is almost equivalent to exit, as both forces discount future yields. The difference is that we assume depreciation starts only after period 2, whereas exit also operates in periods 1 and 2. Thus incorporating an annual exit rate of 0.007 roughly amounts to discounting the annual yield by twice this amount and increasing the depreciation rate by 0.007. The combined impact of these changes is roughly a 2 percent reduction in the returns. We ignore this effect in the main text.

A.1.7 Demand diffusion

Model. We model the demand externality for the special case in which (i) there is no diffusion ($\phi = 0$), and (ii) there is no heterogeneity in λ ($\lambda_m^t = \lambda^0$). This special case is not a bad approximation of reality, and it makes the analysis simple. We assume that the intervention generates an indirect effect in which a share D_m of firms in market m experience demand diffusion. We further assume that demand diffusion is governed by the share of non-competitors in the same market which are treated, and can be written as $D_m = aS_m + v_m$ where v_m captures sampling variation and is independent of all realizations absent the intervention. The assumption that demand diffusion depends directly on the share who are treated, rather than the share who borrow, simplifies the

logic, as otherwise we would need to instrument the latter with the former.

Experiencing demand diffusion means that the firm's perceived quality is shifted by a factor e^{χ} where $\chi \geq 0$. This is compensated for by a shift in the perceived quality of all firms in market m of $e^{-D_m \chi}$ so that to a first-order approximation average quality is unchanged. Let X_i be an indicator for whether firm i experiences the demand externality. Then the perceived quality $\bar{h}_i$ of firm i satisfies $\log \bar{h}_i = \log h_i + \chi X_i - \chi D_m$. We assume that the demand externality is a mistake, so that actual quality is still given by h_i. We use the convention that bar denotes qualities and prices misperceived by the consumer. We do not include bars in the notation for quantities.

Estimating equation. The price index $\bar{P}_m$ for market m satisfies

$$\begin{aligned}\bar{P}_m &= \left(\int_{i \in m} \left(\frac{P_i}{\bar{h}_i} \right)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}} \\ &= \left(\int_{i \in m} \left(\frac{P_i}{h_i} \right)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}} \cdot (D_m e^{\chi(1-D_m)(\sigma-1)} + (1-D_m) e^{-\chi D_m (\sigma-1)})^{\frac{1}{1-\sigma}}.\end{aligned}$$

The log of the second term at $\chi = 0$ is zero, and its derivative with respect to χ is

$$\frac{1}{1-\sigma} \frac{\partial}{\partial \chi} \log(D_m e^{\chi(1-D_m)(\sigma-1)} + (1-D_m) e^{-\chi D_m (\sigma-1)}) = D_m(1-D_m)(\sigma-1) - D_m(1-D_m)(\sigma-1) = 0$$

thus, to a first order approximation in χ, $\log \bar{P}_m \approx \log P_m$.

By (A3) in the proof of Proposition 1, firm revenue is

$$P_i Q_i = \left(\frac{\theta}{\theta - 1} \right) \left(\frac{P_i}{\bar{h}_i} \right)^{1-\sigma} \bar{P}_m^{\sigma-\theta}. \tag{A17}$$

Since $\log \bar{P}_m \approx \log P_m$, substituting in $\bar{h}_t$ implies that in period $t \geq 1$

$$\Delta \log R_i^t \approx (\sigma - 1) \gamma \mu^T \cdot T_i - (\sigma - \theta) \gamma \mu^T \lambda \cdot S_m + (\sigma - 1) \chi \cdot X_i - (\sigma - 1) \chi \cdot D_m$$

and hence revenue is

$$\begin{aligned}\log R_i^t \approx &(\sigma-1)\gamma\mu^T \cdot Post^t \cdot T_i - (\sigma-\theta)\gamma\lambda\mu^T \cdot Post^t \cdot S_m \\ &+ (\sigma-1)\chi \cdot Post^t \cdot X_i - (\sigma-1)\chi \cdot Post^t \cdot D_m + \kappa \cdot Post^t + f_i + \varepsilon_i^t. \end{aligned} \tag{A18}$$

Here ε_i^t is defined exactly as before, and in this special case with no λ heterogeneity only depends on realizations absent the intervention. Thus ε_i^t is orthogonal to $Post^t \cdot T_i$ and $Post^t \cdot S_m$, and is also orthogonal to $Post^t \cdot X_i$ and to $Post^t \cdot D_m$ because X_i is randomly assigned conditional on D_m and $D_m = aS_m + v_m$ with v_m independent of realizations absent the intervention. This equation motivates our empirical approach to estimating the direct and indirect effects induced by demand diffusion.

Welfare. To characterize the welfare effect of demand diffusion, note that due to consumer optimization in the presence of misperception

$$\frac{\bar{h}_i Q_i}{Q_m} = \left(\frac{P_i/\bar{h}_i}{P_m}\right)^{-\sigma} \tag{A19}$$

which implies that

$$(h_i Q_i)^{\frac{\sigma-1}{\sigma}} = \left(\frac{\bar{h}_i}{h_i}\right)^{\frac{(\sigma-1)^2}{\sigma}} Q_m^{\frac{\sigma-1}{\sigma}} \left(\frac{P_i/h_i}{P_m}\right)^{1-\sigma}. \tag{A20}$$

The true (not misperceived) utility from consuming goods in market m is then

$$\begin{aligned} &Q_m^{\frac{\sigma-1}{\sigma}} \int_{i\in m} \left(\frac{\bar{h}_i}{h_i}\right)^{\frac{(\sigma-1)^2}{\sigma}} \left(\frac{P_i/h_i}{P_m}\right)^{1-\sigma} di \\ &\qquad = Q_m^{\frac{\sigma-1}{\sigma}} \int_{i\in m} \left(\frac{P_i/h_i}{P_m}\right)^{1-\sigma} di \cdot (D_m e^{\chi(1-D_m)\frac{(\sigma-1)^2}{\sigma}} + (1-D_m)e^{-\chi D_m \frac{(\sigma-1)^2}{\sigma}}) \end{aligned} \tag{A21}$$

because of the random assignment of demand diffusion. Because of a logic parallel to establishing that $\log \bar{P}_m \approx \log P_m$, the log of the second term, to a first-order approximation, is zero, implying that demand diffusion, to a first order approximation, has no impact on the utility enjoyed from consuming goods in market m. Because the market price index is also approximately the

same as without it, demand diffusion has approximately no impact on spending either. Because it has approximately no impact on the market level utilities and spending, demand diffusion has approximately no impact on the consumer surplus. And because profits in the differentiated markets are proportional to revenue which equals spending on these markets, demand diffusion has approximately no impact on the producer surplus either.

A.2 Evidence: Additional empirical results

Markets. Table A1 reports the distribution of markets across broad product categories, together with basic data about market size and firm size.

Table A1: Distribution of broad product categories across markets

Broad product category	Num of Markets	Num of Sample Firms in Market	Num of Firm Employees
Building Materials	15	45.3	8.4
Furniture	9	25.8	7.5
Cloth and Shoes	12	58.8	8.3
Food	11	31.2	9.5
Electronics	8	36.6	8.6
Vehicle	6	46.5	9.9
Textile	5	51.6	8.5
Daily Necessities	4	32.5	8.2
Toys and Games	3	37.3	7.8
Industrial Park	4	28.8	19.9
Hardware	1	27.0	6.2
Total	78	40.7	9.0

Baseline balance for firms that remain in sample. Table A2 reports balance tests at baseline of firms that remain in the sample up to the 2016 endline survey (columns 1-4) or up to the 2020 follow-up survey (columns 5-8). Similarly to Table 1, these tables report the results of regressing the dependent variable on a constant and on four indicators of different treatment arms: treated firm in a 50% market, untreated firm in a 50% market, treated firm in an 80% market and untreated firm in an 80% market. There are no significant differences across treatment arms, suggesting that selection due to exit or attrition is unlikely to meaningfully affect our estimates.

Baseline balance with respect to direct and indirect effect measures. Table A3 reports balance

Table A2: Baseline balance of treatment arms for firms that remain in sample

	Firms in Sample up to 2016 Endline				*Firms in Sample up to 2020 Follow-up*			
Dep. Var. Measured at Baseline:	Sales	Profit	Num of Employees	Num of Clients	Sales	Profit	Num of Employees	Num of Clients
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated * 50% Market	35.89	-2.302	1.305	-1.962	11.22	-5.390	1.139	-1.894
	(84.73)	(11.96)	(1.241)	(2.048)	(88.83)	(12.98)	(1.096)	(1.700)
Untreated * 50% Market	10.70	-3.050	0.030	0.878	3.096	-5.875	0.325	-0.230
	(56.16)	(8.339)	(1.011)	(2.204)	(77.99)	(10.25)	(1.086)	(2.553)
Treated * 80% Market	3.785	-1.100	0.016	0.416	-0.715	-2.652	-0.061	-0.321
	(51.43)	(7.661)	(0.699)	(1.631)	(70.93)	(10.81)	(0.863)	(1.655)
Untreated * 80% Market	-15.00	-0.494	0.095	2.653	-29.74	-4.453	-0.059	1.916
	(39.33)	(7.115)	(0.662)	(1.980)	(56.32)	(9.801)	(0.812)	(2.015)
Mean of Pure Control	324.56	51.75	8.74	28.62	344.04	54.36	8.88	29.39
Observations	2,482	2,482	2,482	2,482	1,804	1,804	1,804	1,804

Note: Dependent variables are measured at baseline. In columns 1-4 sample is all firms surveyed in the baseline, midline and endline surveys. In columns 5-8 sample is the subset of firms also surveyed in the 2020 follow up. Beyond the regressions shown, in both samples, we regress the four treatment indicators on all baseline variables listed in Table 1 Panels A-D, and find no evidence of joint significance (average p=0.57, lowest p=0.17 across the 8 specifications). Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

tests in which outcomes at baseline are regressed on the treatment and the share of competitors treated. Panel A reports results for all firms in the baseline, Panel B for firms that remained in sample up to the 2016 endline, and Panel C for firms that remained in sample up to the 2020 follow-up survey. We find that our key explanatory variables are balanced with respect to baseline firm characteristics in all three samples.

Effects on loan amount by endline. Table A4 reports direct and indirect treatment effects on the loan amount. The table parallels Table 2 but the outcome variables are loan amounts measured in 10,000 RMBs. Columns 1-3 show that the treatment had a significant and large direct and indirect effect on the amount borrowed through the new program. Columns 4-5 show impacts on the amount borrowed from other sources and indicate that the new borrowing did not crowd out other borrowing: the coefficients are positive, and marginally significant in one case, suggesting that the new loan may have even encouraged other borrowing. Columns 6-7 show impacts on the total loan amount. Here the effects are slightly larger than but not statistically different from those in columns 2-3, and confirm that the loan program generated new borrowing both directly and

Table A3: Baselne balance of explanatory variables

Dep. Var. Measured at Baseline:	log Sales	Profit (10,000 RMB)	log Num of Employees	log Wage Bill	Fixed Assets (10,000 RMB)	log Material Cost
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Full sample						
Treated	-0.008	1.413	0.003	0.006	1.274	-0.027
	(0.041)	(4.047)	(0.038)	(0.040)	(2.552)	(0.051)
Share Competitors Treated	-0.024	-3.885	-0.031	-0.032	0.181	0.012
	(0.052)	(5.123)	(0.048)	(0.051)	(3.302)	(0.064)
Mean of Pure Control	5.453	51.945	1.956	2.675	12.005	4.925
Observations	3,173	3,173	3,173	3,167	3,173	3,173
Panel B. Firms in sample up to 2016						
Treated	-0.00565	2.069	0.0121	0.00399	1.894	-0.0358
	(0.0456)	(3.865)	(0.0420)	(0.0443)	(2.992)	(0.0572)
Share Competitors Treated	-0.0162	-5.147	-0.0145	-0.00929	-2.011	0.0405
	(0.0573)	(4.804)	(0.0528)	(0.0554)	(3.767)	(0.0711)
Mean of Pure Control	5.469	51.753	1.948	2.670	12.081	4.945
Observations	2,482	2,482	2,482	2,480	2,482	2,482
Panel C. Firms in sample up to 2020						
Treated	-0.00470	3.202	0.0220	0.00846	2.094	-0.0416
	(0.0541)	(3.999)	(0.0461)	(0.0494)	(3.749)	(0.0683)
Share Competitors Treated	-0.0411	-8.542	-0.0411	-0.0251	-1.674	0.0214
	(0.0683)	(5.131)	(0.0597)	(0.0636)	(4.789)	(0.0857)
Mean of Pure Control	5.485	54.355	1.962	2.681	13.886	4.958
Observations	1,804	1,804	1,804	1,802	1,804	1,804

Note: Dependent variables measured at baseline. Sample in Panel A is all firms we surveyed in the baseline survey; in Panel B all firms we surveyed in the baseline, midline and endline surveys; in Panel C all firms we also surveyed in the 2020 follow up. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

indirectly. Note that the sample size in columns 1-3 is larger than in columns 4-7, because we have data on borrowing through the new product for all firms (from the bank), but on other borrowing only for firms in the endline survey. Rerunning columns 2-3 for that subsample yields results even more similar to columns 6-7.

Alternative specifications. Table A5 reports alternative specifications for our main effects. In columns 1-3 we regress three main outcomes (log sales, profit and log employment) directly on the treatment arms. Consistent with the direct effect, relative to untreated firms in treated markets,

Table A4: Effect of treatment on loan amount by endline

Dep. Var.:	Amount Borrowed with New Loan Product			Amount Borrowed from Other Sources		Amount Borrowed from Any Source	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	8.079***	8.929***		1.841*		12.37***	
	(0.955)	(0.966)		(0.995)		(1.756)	
Untreated * Share of Peers Treated		4.323***		0.673		5.811***	
		(0.911)		(1.247)		(1.517)	
Treated * 50% Market			8.908***		2.024		12.84***
			(1.494)		(1.272)		(2.103)
Treated * 80% Market			8.935***		1.786		12.27***
			(1.096)		(1.119)		(2.000)
Untreated * 50% Market			2.405*		0.284		3.154**
			(1.338)		(1.002)		(1.541)
Untreated * 80% Market			3.456***		0.488		4.620***
			(0.764)		(1.171)		(1.386)
Constant	1.571***	0.720**	0.719**	8.558***	8.577***	9.438***	9.437***
	(0.340)	(0.305)	(0.302)	(0.476)	(0.482)	(0.635)	(0.637)
Mean of Pure Control	0.719	0.719	0.719	8.577	8.577	9.437	9.437
Observations	3173	3173	3173	2,668	2,668	2,668	2,668

Note: Dependent variable, measured in 10,000 RMBs, is the amount borrowed by endline, in columns 1 to 3 from the new loan program; in columns 4 and 5 from other sources; in columns 6 and 7 from all sources. Sample in columns 1-3 is all firms, and in columns 4-7 is firms in the endline survey. Share of peers treated is the share of other firms in the market which are treated. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

treated firms in treated markets experienced higher growth. Consistent with the business stealing effect, relative to pure control firms, untreated firms in a treated markets experienced lower growth. These results align closely with our findings in Table 3. In columns 4-6 we estimate for the three main outcomes an ANCOVA version of our main specification (McKenzie 2012) and find that the results are very similar to those in Table 3.

Levels and IHS, including firms that shut down. In Table A6 we study robustness to alternative functional forms and to including firms after they shut down, imputing zero values for their performance measures. In Panel A we measure sales, profit and employment in levels. Columns 1-3 report regressions in the main sample used in Table 3 in which we do not include observations

Table A5: Alternative specifications

A. Effect of Treatment Arms				*B. ANCOVA*			
Dep. Var.:	log Sales	Profit (10,000 RMB)	log Num of Employees	Dep. Var.:	log Sales	Profit (10,000 RMB)	log Num of Employees
	(1)	(2)	(3)		(4)	(5)	(6)
Post*50% Treated Market*Treated	0.0985	12.97***	0.0719	Treated	0.099***	12.66***	0.078***
	(0.0982)	(3.321)	(0.0600)		(0.035)	(2.724)	(0.026)
Post*80% Treated Market*Treated	0.0984**	13.20***	0.0719*	Share Competitors Treated	-0.088**	-10.56**	-0.072**
	(0.0436)	(3.069)	(0.0371)		(0.042)	(4.652)	(0.036)
Post*50% Treated Market	-0.0640**	-9.168***	-0.0476	Baseline Value of Outcome	0.801***	0.580***	0.838***
	(0.0313)	(2.917)	(0.0497)		(0.019)	(0.063)	(0.015)
Post*80% Treated Market	-0.0660	-8.117**	-0.0482				
	(0.0430)	(3.975)	(0.0412)				
Firm FE and Post	Yes	Yes	Yes	Firm FE and Post	No	No	No
Mean of Pure Control	5.45	51.95	1.96	Mean of Pure Control	5.45	51.95	1.96
Observations	8612	8612	8612	Observations	5439	5439	5439

Note: Columns 1-3 report impacts on main outcomes of the treatment arms, 4-6 report impacts on main outcomes of the direct and indirect effects in an ANCOVA. Mean of pure control computed at baseline. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

for firms after they shut down. The qualitative results are the same as in Table 3. Columns 4-6 report regressions in the larger sample in which we impute zero for sales, profit and employment for firms after they shut down. The results are similar, and statistically indistinguishable, from those in columns 1-3. In Panel B we measure sales, profit and employment with the inverse hyperbolic sign. Here too, we find that the results are qualitatively the same as in Table 3 in the main sample(columns 1-3), and are statistically unchanged when we include firms with zero imputed outcomes after they shut down (columns 4-6).

Heterogeneity by treatment status and time. In Table A7 we estimate key heterogeneities. Columns 1-3 report results from a "fully saturated" specification in which we include the interaction of Share Peers Treated with both Treated and Untreated. The coefficients are insignificant except for the negative indirect effect on untreated firms. However, the magnitudes are similar to those in Table 3 and we cannot reject the model-implied restriction that indirect effects are identical for treated and untreated firms. Columns 4-6 report impacts separately at midline and at endline. The effects at endline are slightly larger and more significant than the effects at midline, suggesting that the impacts do not fade over time.

Table A6: Levels and IHS, including firms that exit

Panel A. Levels	*Main Sample*			*Including Firms that Exit*		
Dep. Var.:	Sales (10,000 RMB)	Profit (10,000 RMB)	Num of Employees	Sales (10,000 RMB)	Profit (10,000 RMB)	Num of Employees
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Treated	65.37***	12.64***	0.744**	70.65***	11.61***	0.819**
	(13.62)	(3.099)	(0.317)	(14.26)	(2.986)	(0.312)
Post*Share Competitors Treated	-48.93***	-9.478*	-0.854*	-52.92***	-7.860*	-0.882*
	(17.36)	(4.802)	(0.492)	(16.86)	(4.670)	(0.469)
Firm FE and Post	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Pure Control	323.71	51.95	8.82	323.71	51.95	8.82
Observations	8,612	8,612	8,612	8,799	8,799	8,799
Panel B. IHS	*Main Sample*			*Including Firms that Exit*		
Dep. var.:	*asinh* (Sales)	*asinh* (Profit)	*asinh* (Num Employees)	*asinh* (Sales)	*asinh* (Profit)	*asinh* (Num Employees)
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Treated	0.0993***	0.300***	0.0907**	0.212**	0.209**	0.108***
	(0.0350)	(0.0822)	(0.0348)	(0.0836)	(0.0992)	(0.0338)
Post*Share Competitors Treated	-0.0855**	-0.273**	-0.0988***	-0.164*	-0.134	-0.0979**
	(0.0410)	(0.123)	(0.0365)	(0.0972)	(0.141)	(0.0400)
Firm FE and Post	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Pure Control	6.14	4.15	2.67	6.14	4.15	2.67
Observations	8,612	8,612	8,612	8,799	8,799	8,799

Note: Panel A reports impacts on level sales, profit and employment, Panel B on the inverse hyperbolic sine of sales, profit, and employment. Sales and profit always measured in units of 10,000 RMB. In Columns 1-3 the sample is the main sample used in Table 3. In Columns 4-6, we keep in the sample firms that exit and impute values of zero for sales, profit and employment. Mean of pure control computed at baseline. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

Other business outcomes. Table A8 reports our main reduced form specification with other business practices as the outcome variables. The results are discussed in Section 4.3.

Alternative explanations. Table A9 reports impacts on outcomes that help rule out some alternative explanations. Columns 1-3 are about experimenter demand. In the endline survey, we collected data about sales directly from the firms' books. Column 1 shows impacts on this measure of sales. Although less significant, the estimates are similar to those in Table 3. In column 2 we also include the self-reported value of sales in the baseline, which was collected before the intervention and hence is plausibly immune to experimenter demand. We find significant effects that are similar

Table A7: Main effects by treatment status and over time

A. Indirect Effect by Treatment Status				*B. Effects over Time*			
Dep. Var.:	log Sales	Profit (10,000 RMB)	log Num of Employees	Dep. Var.:	log Sales	Profit (10,000 RMB)	log Num of Employees
	(1)	(2)	(3)		(4)	(5)	(6)
Post*Treated	0.070	7.729	0.032	Midline*Treated	0.0951**	11.81**	0.0830**
	(0.087)	(11.47)	(0.054)		(0.0426)	(5.491)	(0.0372)
Post*Share Competitors Treated*Treated	-0.049	-3.398	-0.013	Midline*Share Competitors Treated	-0.0770	-9.668	-0.0732
	(0.106)	(14.46)	(0.064)		(0.0465)	(6.397)	(0.0492)
Post*Share Competitors Treated*Untreated	-0.094**	-10.90**	-0.078*	Endline*Treated	0.103***	13.43***	0.0664***
	(0.045)	(4.230)	(0.044)		(0.0378)	(3.180)	(0.0248)
				Endline*Share Competitors Treated	-0.0935**	-9.306**	-0.0591*
					(0.0459)	(4.334)	(0.0328)
Firm FE and Post	Yes	Yes	Yes	Firm FE and Mid/Endline	Yes	Yes	Yes
Mean of Pure Control	5.45	51.95	1.96	Mean of Pure Control	5.45	51.95	1.96
Observations	8,612	8,612	8,612	Observations	8,612	8,612	8,612

Note: Mean of pure control computed at baseline. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Direct and indirect effects: Other business outcomes

Dep. var.:	Trade Credit Supplier	Trade Credit Client	log Markup	log Rent	log Num of Suppliers	Advertisement (10,000 RMB)
	(1)	(2)	(3)	(4)	(5)	(6)
Post*Treated	0.068***	0.075***	0.015	0.048	0.051	0.210*
	(0.014)	(0.020)	(0.023)	(0.099)	(0.034)	(0.106)
Post*Share Competitors Treated	-0.041**	-0.072***	-0.040	0.011	-0.041	0.037
	(0.018)	(0.023)	(0.026)	(0.162)	(0.035)	(0.075)
Firm FE and Post	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Pure Control	0.174	0.348	0.389	1.634	1.216	0.366
Observations	8,222	8,221	8,612	8,220	8,716	8,220

Note: In column 3 markup is defined as the ratio of revenue to cost, the latter measured as total wage and material spending. Mean of pure control computed at baseline. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

to those in Table 3. In column 3 we report impacts on the difference between the self reported and the book value (in the endline) and find essentially null effects.

Column 4 reports impacts on an index of managers' financial knowledge. This measure was constructed by standardizing the average of answers to the following binary questions: (i) whether

Table A9: Alternative explanations

Dep. Var.:	Log Book Sales	Log Book Sales Endline, Log Reported Baseline	Log Reported - log Book Sales	Financial Knowledge	Difficulty of Getting Loans
	(1)	(2)	(3)	(4)	(5)
Treated*Post	0.106*	0.110**	0.0128	0.096	-0.075
	(0.0589)	(0.045)	(0.0835)	(0.065)	(0.075)
Share of Competitors Treated*Post	-0.0989	-0.091*	-0.0222	-0.035	-0.024
	(0.0794)	(0.050)	(0.112)	(0.071)	(0.082)
Firm FE and Post	No	Yes	No	No	No
Mean of Pure Control	5.51	5.454	-0.033	-0.003	0.026
Observations	2,530	5,703	2,530	2,714	2,714

Note: In columns 1, 3-5 sample is endline only; in column 2 baseline and endline. Book sales is the value of sales that enumerators took from the firm's book. In column 2, outcome is log of book sales in the endline and log of reported sales in the baseline. Financial knowledge is standardized average of answers to 5 questions described in the Appendix about manager's financial knowledge. Difficulty of getting loans is standardized answer to manager's evaluation of perceived difficulty of obtaining a loan from a formal bank. Mean of pure control is computed for the first wave in which outcome is measured. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

our partner bank offers loans to micro and small enterprises (yes is correct); (ii) whether the interest rate charged by formal lenders is higher than that of moneylenders (no is correct); (iii) whether the manager knows of financial products they can invest in (yes means higher knowledge); (iv) whether the manager knows that investment products involve risks (yes means higher knowledge); and (v) whether the manager reads the relevant written material when handling their finances (yes means higher knowledge). We estimate small and insignificant effects.

Column 5 reports impacts on the difficulty of getting a loan, measured by standardizing the response to a question we asked in the endline about how challenging the manager finds, on a five-point scale, to get loans from a formal bank. We estimate small and insignificant effects.

Demand diffusion: additional results. Table A10 reports additional results on the demand diffusion effect. Columns 1-4 show that the share of local non-competitors treated does not significantly increase consumer satisfaction, providing further evidence that the effect is not driven by the diffusion of borrowing.

Columns 5-7 are an empirical version of equation (A18). Here we ask whether demand diffusion

Table A10: Demand diffusion: Additional results

Dep. Var.:	Service Quality	Shopping Environment	Value for Money	Overall Satisfaction	log Sales	Profit	log Num of Employees
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	0.738***	0.987***	0.616***	0.839***			
	(0.100)	(0.101)	(0.084)	(0.068)			
Share Local Competitors Treated	-0.288*	-0.253**	-0.173*	-0.169**			
	(0.133)	(0.105)	(0.097)	(0.081)			
Share Local Non-Competitors Treated	0.131	0.020	0.033	0.048			
	(0.142)	(0.108)	(0.084)	(0.087)			
Post*Treated					0.095**	10.14***	0.079**
					(0.040)	(3.007)	(0.031)
Post* Share Local Competitors Treated					-0.122**	-12.42**	-0.061
					(0.053)	(5.852)	(0.039)
Post*Share Local Non-Competitors Treated					0.190***	14.32***	0.066**
					(0.050)	(5.234)	(0.030)
Post*Share Local Competitors with High Share of Local Non-competitors Treated					-0.04	-10.55**	-0.089**
					(0.088)	(5.299)	(0.043)
Firm FE and Post	No	No	No	No	Yes	Yes	Yes
Mean of Pure Control	-0.004	0.043	0.024	0.021	5.45	51.95	1.96
Observations	1,804	1,804	1,804	1,804	8,220	8,220	8,220

Note: Columns 1-4 use the 2020 follow-up data and include the share of non-local competitors and non-local non-competitors treated. Columns 5-7 use the full panel for firms in the endline, and include Post interacted with the share of non-local competitors and non-local non-competitors treated. The share of local competitors with high share of local non-competitors treated is defined with the above-median share of local non-competitors treated. Indicators for no peers in the relevant group, interacted with Post, are included. In columns 1-4, we include controls for baseline employment, and for the county and broad product category of the market. Mean of pure control computed in first wave in which outcome is measured. Standard errors clustered at the market level. *** p<0.01, ** p<0.05, * p<0.1.

generates business stealing: intuitively, whether a firm experiences a greater reduction in revenue if more of its competitors get demand diffusion, which in turn results from their non-competitor neighbors getting treated. This is a "second-degree" indirect effect. To estimate it, we include the interaction of *Post* with the share of local competitors being exposed to demand diffusion. And we measure the latter with an indicator for the firm having a higher-than-median share of local non-competitors treated. We find a significant negative effect for two of the three outcomes, profit and the log number of employees, providing evidence consistent with demand diffusion generating business stealing effects. But because the new variable is correlated with the share of competitors treated (which measures the business stealing effect of the treatment), we interpret these results as

suggestive.

Welfare and return to capital: different elasticities. To show the sensitivity of our welfare results to the elasticity of substitution we replicate Tables 10 and 11 using $\sigma = 4$ and $\sigma = 8$. Table A11 reports the welfare estimates. The impact on the producer surplus is unchanged because it is computed directly from the regression coefficients. The impact on the consumer surplus, as well as the total effect, is larger for $\sigma = 4$, and smaller for $\sigma = 8$, but the qualitative patterns are similar to the results in the main text. Table A12 reports the estimated returns to capital. Here too, the results are qualitatively similar to those in the main text.

Table A11: Robustness of welfare gain estimates

Welfare Gain per	Treat All Firms		Treat 50% of Firms	
Firm in Market	Share of Profit (%)	USD	Share of Profit (%)	USD
Panel A: Sigma=4				
Producer Surplus	4.1	3,566	2.0	1,778
	(4.4)	(3,904)	(2.2)	(1,952)
	[-5, 12]	[-4,174, 10,919]	[-2, 6]	[-2,210, 5,310]
Consumer Surplus	21.1	18,561	10.5	9,281
	(7.6)	(6,703)	(3.8)	(3352)
	[7, 37]	[6,551, 32,694]	[4, 19]	[3,276, 16,343]
Spillover			3.6	3,135
			(1.8)	(1,598)
			[1, 8]	[713, 7,229]
Total	25.1	22,119	16.1	14,203
	(9.9)	(8,741)	(6.7)	(5,859)
	[7, 46]	[5,996, 40,627]	[4, 30]	[3,689, 26,909]
Panel B: Sigma=8				
Producer Surplus	4.1	3,566	2.0	1,778
	(4.4)	(3,904)	(2.2)	(1,952)
	[-5, 12]	[-4,174, 10,919]	[-2, 6]	[-2,210, 5,310]
Consumer Surplus	9.0	7,951	4.5	3,980
	(3.3)	(2,873)	(1.6)	(1,437)
	[3, 16]	[2,809, 14,009]	[2, 8]	[1,400, 7,009]
Spillover			1.9	1,629
			(1.1)	(961)
			[0, 5]	[123, 3,980]
Total	13.1	11,517	8.4	7,396
	(6.3)	(5,518)	(4.2)	(3,666)
	[1, 25]	[634, 22,357]	[0, 17]	[423, 14,978]

Note: Bootstrapped standard errors in round brackets, and bootstrapped bias-corrected percentile confidence intervals in square brackets, are computed by bootstrapping our estimation procedure 1,000 times, drawing markets with replacement. USD values are calculated using the average annual exchange rate rate during midline and endline (6.465).

Table A12: Robustness of return to capital decomposition

	Sigma=4	Sigma=8
Private Return (%)	74.2	74.2
	(12.9)	(12.9)
	[46, 98]	[46, 98]
Business Stealing (pp)	-56.3	-56.3
	(23.4)	(23.4)
	[-104, -13]	[-104, -13]
Consumer Surplus (pp)	62.8	31.7
	(19)	(10.7)
	[26, 102]	[12, 55]
Social Return (%)	80.8	49.6
	(25.3)	(20.6)
	[29, 127]	[5, 86]

Note: Bootstrapped standard errors in round brackets, and bootstrapped bias-corrected percentile confidence intervals in square brackets, are computed by bootstrapping our estimation procedure 1,000 times, drawing markets with replacement. In draws with negative raw yields (about 10% of cases), we approximate the internal rate of return with the yield.

B Script and Survey

Bank Agent Visiting Script (Treated firms only)

Hello Sir/madam! My name is _____, I'm a loan agent from *** bank. We are visiting you today to introduce a new loan program that we designed for businesses like yours. We have heard from many business owners that it is very difficult for you to get loans from formal sources, because of the requirement of collateral. In this new program, we do not require any collateral. If you are interested in applying, we will collect information about your business and your own background information through asking you to fill out application forms and requesting for information from the market office. To save your time, we will assist you in the application process and help you to fill out the application forms if you can provide us with all the required documents and information. The monthly interest rate is 0.7%, and once your loan application is approved, we will assign you a credit line of up to 30% of the value of your net assets, not exceeding 500,000 RMB. Then you can borrow any amount up to your assigned credit line. You will need to make monthly interest payments and repay the loan within two years.

If you have any questions about the new loan program now or at any future date, you can contact us at this phone number ***.

FIRM SURVEY FOR CREDIT ACCESS EXPERIMENT IN CHINA

We provide below the English version of all questions used in our surveys. For presentational clarity in a few places below we changed the organization of the questions into sections relative to that used in the data collection. Four survey waves were conducted: baseline (2013 summer), midline (2015 summer), endline (2016 summer), and a short long-term follow-up (2020 summer). Because not all questions were included in all waves, we clarify below which waves contained which questions.

A – CONTROL INFORMATION – To be completed before the interview

County/District: _____
Town/Street: _____
Firm address: _____
Firm name：_____
Firm ID: _____
Name of firm owner：_____Cellphone：_____
Position/job title of the INTERVIEWEE within firm：_____（1-Owner 2-Spouse of owner 3-Parents of owner 4-Children of owner 5-Manager 6-Worker 7-Other（Specify_____））
Name of the INTERVIEWEE：_____ Cellphone：_____
Name of the INTERVIEWER：_____
DATE of the interview：_____

MAIN SURVEY QUESTIONS

In the 2020 follow-up we only asked those questions for which we explicitly indicate this below.

B – GENERAL INFORMATION

1. Year the company was established (USE FOUR DIGITS): _____
2. 1) Registration type (1=enterprise 2=individual business): _____ If answer is 1, skip to Q3.
   2) Do you plan to change registration to "Enterprises" this or next year (1=yes 0=no): _____
3. What's the industry that this firm belong to (1=manufacturing 2=retail 3=service) _____

4. What is this firm's current legal status? (1=State-owned 2=Shareholding 3=Sole proprietorship 4=Partnership 5=Limited partnership 6=Foreign 7=Others – Specify) _____
5. What is your ownership share in the company? _____
6. Are you the largest owner? (1=yes 0=no)_____

7. Please list the three main businesses, ranked by the share in total sales:
   1) Main business 1 _____Code_____ Share in total sales _____ (%)
   2) Main business 2 _____Code_____ Share in total sales _____ (%)
   3) Main business 3 _____Code_____ Share in total sales _____ (%)

8. 2020 Follow-up only:
   1) What is the current price at which you sell the top product (main bus 1)_____
   2) What was the price at which you sold the top product (main bus 1) in 2016 _____

9. From June last year until now, what's the number of new products produced or sold _____
10. From June last year until now, did you renovate your factory or stores? (1=yes 0=no)_____

11. Labor:

	Number	Relatives（%）	Female（%）	Local（%）	***2020 follow-up only:*** *% Education in 2016*				Monthly Salary（RMB）	Number of months work in a year
					>=Master	College	High school	<=Middle School		
Full time										
Part time										

C. MANAGER INFORMATION

1. Gender (1=male 2=female) _____
2. Year of birth _____
3. Highest level of education (00=below primary school 01-parimary school 02-secondary school 03-high school 04-college 05-master 06-doctor)_____
4. Are you a member of the Communist Party of China (1=yes 0=no)_____
5. Are you a member of renda or zhengxie (1=yes 0=no)_____
6. Do you have working experience in government or state-owned enterprises (1=yes 0=no)_____

D. BUSINESS ENVIRONMENT

1. Please describe the impact of the following factors on the growth of your firm (1=big impact 2=moderate impact 3=small impact 4=no impact): ___ financing; ____hiring; ____market; _____technology; ____local government support; ____other government support; ____property right protection; ____infrastructure
2. Is it difficult to get the operation license? (1=very difficult 2=moderately difficult 3=difficult 4=not very difficult 5=not difficult at all) _____
3. Is it difficult to get a loan from formal banks? (1=very difficult 2=moderately difficult 3=difficult 4=not very difficult 5=not difficult at all) _____
4. Is it difficult to buy/rent the factory/store? (1=very difficult 2=moderately difficult 3=difficult 4=not very difficult 5=not difficult at all) _____
5. How much time do you spend on communicating with government officials? (1=a lot 2=moderate 3=average 4=not very much 5=not at all) _____
6. Are you enjoying any favorable tax policies? (1=yes, 0=no) _____
7. How many business partners does your company have? _____
8. How many clients does your company have? _____
9. How many suppliers does your company have? _____
10. ***Endline only:*** Did you change your main suppliers from June last year until now? (1=yes, 0=no) _____

E. FINANCING

1. During the last fiscal year how much did this firm receive in subsidies and grants from the national or regional government and from other governmental sources? _____ (10,000 RMB)

2. Please tell us about each loan your have used to finance operations or investments in this company which you haven't yet paid back.

	Type of source	Original loan amount (RMB)	Amount outstanding today (RMB)	Average monthly interest (in percentage points)	Year and month of borrowing	Year and month of expiration	***Endline only:*** Usage of loan	Type of loan	Type of collateral	Frequency of payments
1										
2										
3										
4										
5										
CODES	1 partner bank 2 other banks 3 relatives /friends 4 money lender 5 business partners 6 government 7 other-specify						1 inventory 2 renovation 3 rent 4 hiring 5 new project 6 other-specify	1 credit 2 collateral 3 guarantor 4 subsidized 5 other-specify	1 land 2 building 3 machine 4 other 5 none	1 monthly 2 quarterly 3 annually 4 pay on the expiration date 5 other-specify

3. What's the share of credit demand that you have managed to borrow? _____%
4. If no bank loan, have you applied for one (1=yes, 0=no) _____If no, skip to 6.
5. If yes, what are the reason of rejection (1=no collateral 2=no guarantor 3=credit line not high enough 4= others)_____
6. Why didn't you apply for bank loan? (1= No need for a loan 2= Interest rates are too high 3= Application procedures for loans or lines of credit are complex 4=application cost too high (time, connection, etc.) 5= Collateral requirements are too high 6= Size of loan and maturity are insufficient) _____
7. In addition to the loan, how much is the other accounts payable? _____

8. ***Endline only:*** Financial knowledge test
   1) True or false: The study bank does not offer loans to micro and small enterprises _____
   2) True or false: The interest rate charged by formal banks is usually much higher than that by money lenders _____
   3) Do you know other than saving, there are other financial products you can invest on (1=yes, 0=no) _____
   4) When you buy investment products, what's your understanding regarding the expected risk and returns? _____ (A. Financial investment products has risks as other investments; B. Similar as saving products, financial investment products do not have any risks.)
   5) When you handle financial business, do you read the contracts, instructions, and other relevant written materials provided by financial institutions? (1=yes, 0=no) _____

F. ACCOUNTING (June last year until now)

1 Sales	
2 Average sales growth in the past 3 years（%）	
3 Total operation cost	
3a Inputs and materials	
3b Wage	
3c Rent	
3d Utility cost	
3e Interest payment	
3f Advertisement cost	
3g Travel, meals, office suppliers	
4 Total profit	
5 Tax	
6 Net profit	
7 Current assets	
8 Fixed assets	

G. MANAGEMENT

1. Endline only: Inventory

1) How did you keep track of the inventory of your business? (1= by hand; 2=on the computer; 3=Don't keep track; 4=other (specify): _____)_____
2) Did you have a designated area for inventory storage (1=yes, 0=no) _____
3) How often did you buy inventory? Once every _____ months

2. Trade credit:

1) Do any of your suppliers allow you to pay for the materials more than one week after the delivery(1=yes, 0=no) _____ If no, go to 2.c.
2) What percent of your inputs or materials do you pay your suppliers at least one week after delivery_____
3) Do you allow your clients to pay you for your products more than one week after delivery? _____ If no, go to 3.
4) What percent of your products do you allow your clients to pay you after delivery_____

3. Endline only: book sales

1) Do you have an accounting book? (1=yes, 0=no) _____
2) Can you show us the book value of sales (10,000RMB) _____

H. PRESENT BIAS, RISK PREFERENCE, AND CONFIDENCE

1. Imagine that you have just won 10,000 RMB in a lottery. You can take the 10,000 RMB today. However, if you are willing to accept wait for one or two months, you will get more than 10,000 RMB. I am going to describe some situations and you tell me which you would accept.
   Surveyor: start with the first option and only continue with the following if they reject the option you offer.

Option 1	Option 2	Choice	Option 1	Option 2	Choice
Take10000 today	Take 11,000 after one month		Take10000 today	Take 11,000 after two month	
Take10000 today	Take 12,000 after one month		Take10000 today	Take 12,000 after two month	
Take10000 today	Take 13,000 after one month		Take10000 today	Take 13,000 after two month	
Take10000 today	Take 14,000 after one month		Take10000 today	Take 14,000 after two month	
Take10000 today	Take 15,000 after one month		Take10000 today	Take 15,000 after two month	
Take10000 today	Take 16,000 after one month		Take10000 today	Take 16,000 after two month	
Take10000 today	Take 17,000 after one month		Take10000 today	Take 17,000 after two month	

2. Now we ask you to make some hypothetical risky decisions. Each decision is a paired choice between Option 1 and Option 2. Options themselves are lotteries in which you can gain varying amounts of money. The outcomes of these lotteries are determined by the roll of a ten-sided die.

Option 1	Option 2	Choice
RMB 1800	10% of RMB 3850, 90% of RMB 100	
RMB 1800	20% of RMB 3850, 80% of RMB 100	
RMB 1800	30% of RMB 3850, 70% of RMB 100	
RMB 1800	40% of RMB 3850, 60% of RMB 100	
RMB 1800	50% of RMB 3850, 50% of RMB 100	
RMB 1800	60% of RMB 3850, 40% of RMB 100	
RMB 1800	70% of RMB 3850, 30% of RMB 100	
RMB 1800	80% of RMB 3850, 20% of RMB 100	
RMB 1800	90% of RMB 3850, 10% of RMB 100	

3. Please estimate the probability of the following situation (1=for sure 2=very likely 3=likely 4=not very likely 5=impossible 99=don't know):
   1） Your profit in the last fiscal year is in the top half of the profits of firms in your industry and in this province that employ approximately the same number of workers_____
   2） Your profit growth between the last two fiscal years is in the top half of the profit growth rates between the last two fiscal years of firms in your industry and in this province that employ approximately the same number of workers_____
   3） The profits of your company are going to increase between this and the last fiscal year_____
   4） The profits of your company are going to increase by more than 10% between this and the last fiscal year_____
   5） The total output of your industry in your province in this calendar year will decrease_____
   6） The GDP growth in China in this calendar year will decrease_____

I. *2020 FOLLOW-UP ONLY:* CONSUMER SATISFACTION SURVEY

1. Are you a regular customer in this store? (1=yes, 0=no) _____
2. Did you receive valuable advice from sellers? (1=yes, 0=no) _____

3. On a scale of 1-5 with 1 being the worst and 5 the best, how would you rank this store in terms of (i) product quality_____ (ii) service quality_____ (iii) shopping environment_____ (iv) value for money_____ (v) overall satisfaction_____?