I. Content

1. We address concerns arising from the difference in sample sizes between the treatment and control group. We demonstrate that both the difference in sample sizes, as well as the hump in the distribution of firms above the 200K threshold are consistent with the mechanism described in the paper.

2. We show that narrowing the range of the analysis sample results in similar point estimates, but the statistical significance of some results becomes marginal (Table IA I).

II. Difference in Sample Size Between Treatment and Control Group

A. Concern for Empirical Strategy

In the paper we show that the number of treatment firms is one fourth the number of control firms (Table I). One possibility is that this difference may be due to partial backfilling of information for borrowers below the $200K threshold as a result of the registry expansion. To the best of our knowledge backfilling was not incomplete. Since we cannot rule this out completely we set out why this possibility is not a concern for our empirical results in this appendix. The possibility of partial backfilling raises a concern for our empirical strategy if this selective backfilling was related to (negative) future credit outcomes. This could produce differential outcomes for each group. That is potentially a concern for the results from specification (5) in Table II.

This highlights the importance of specification (6) in the paper and the associated results in Table III. There identification comes from comparing the lending response of banks that assigned a rating of 2 to firms in the treatment group to the lending response of banks that
assigned a 1 to the very same set of firms. As we explain in the paper, specification (6) contains “an interaction between the treatment dummy and time dummies, which account for aggregate shocks that affect differentially borrowing by treatment and control firms. If there is any selection that makes treatment and control firms different, for example, from selective backfilling of firms below the $200,000 threshold, this control would absorb that difference.” Since these results come from identification within the same set of firms they cannot be explained by a selection of which firms have their data backfilled. Moreover, since the results in Table 3 line up very closely with those in Table 2 this indicates that the results there are unlikely to be affected by selection either.

B. Difference in Sample Sizes: Discussion

We provide a possible explanation for the difference in the size of the treatment and control groups. Our primary focus is to show that the difference could arise due the debt dynamics introduced by information sharing, regardless of whether backfilling of firm information is complete or incomplete.

To start, the difference in the number of treatment and control firms is not due to a jump in the population of firms around the $200K debt threshold. Figure IA 1, Panel A, shows the population distribution in March 1998. There is an increase in the population of borrowers to the right of the $200K threshold, but it is not of the order of magnitude that the one observed in the sample. This implies that the striking difference in the number of treatment and control firms is primarily a feature of the relative number of firms who satisfy the specific conditions required to be included in our sample. That is, firms need to be categorized as treatment (control) firms: January to April debt between $150K and $200K ($200K and $250K) and all ratings during this period no worse than 2.

The second observation is that the hump in the distribution of borrowers around the $200K threshold in the population (and the sharp jump in the sample) occur most saliently in 1998. As a possible counterfactual, we draw the firm population distribution in March 1999
in Figure IA 1, Panel B. The hump in the distribution is mostly smoothed out a year after the beginning of our sample. Figure IA 2 shows the distribution of the treatment and control firms by total debt if the sample is constructed using the 1999 data (the plot parallels Figure 1 in the paper). The relative number of firms in the treatment and control firms is different—there is a 200 firm jump in the distribution of firms right above and right below the 200K threshold—but not as sizeable as in 1998.

We now demonstrate that both the difference in sample sizes between the treatment and control groups, as well as the hump in the distribution of the general population of firms above the 200K threshold are consistent with the mechanism described in the paper.

C. Debt Growth Dynamics and Firm Distribution

The difference in the size of treatment and control groups in March of 1998 can potentially be explained by the fact that information sharing reduces the growth rate of a firm’s debt (as we show in Table VI Column 1). We show this using a simple simulation of firm debt growth.

Suppose that firm $i$’s debt in month $t$ grows in the following way

$$D_i^t = (1 + g (I_i^t) + e_i^t) D_{i-1}^t.$$  \hfill (IA 1)

The variable $I_i^t$ is a binary variable that indicates whether or not firm $i$’s debt has satisfied the conditions for information sharing. To simulate the distribution of borrowers in March 1998 we use the same rule that was applied in the registry prior to the expansion: debt must have passed above $200K in the previous 12 months for a firm’s information to be shared. As our empirical work indicates, information sharing can lower the rate at which a firm’s debt grows so we assume

$$g (I_i^t = 1) < g (I_i^t = 0).$$  \hfill (IA 2)

The variable $e_i^t$ captures idiosyncratic noise in a firm’s debt growth. For simplicity we assume
this is mean zero and uniformly distributed on the interval \([-a, a]\).

We numerically simulate the debt growth distribution by supposing that in January 1997 there is a population of 200,000 borrowers with debt uniformly distributed between 0 and $1,000,000. Each month debt growth evolves according to (IA 1) using the following parameters: \( g(I_i^t = 0) = 5\% \), \( g(I_i^t = 1) = 1\% \), and \( a = \frac{1}{24} \). We present the results from averaging 20 simulations to eliminate noise. The simulated distribution of borrowers in March 1998 around $200,000 is shown in Figure IA 3 Panel A. This shows a clear concentration of borrowers above the $200,000 threshold. This comes from a congestion effect as debt growth slows as firms pass over $200,000. This can explain why in 1998 the actual distribution of borrowers in the registry exhibits a hump around this point (Figure IA 1 Panel A).

Next, we create a treatment and control sample using the exact rules we applied in our paper: a firm’s debt must be between $150K and $200K ($200K and $250K) from January to March 1998 to be included in the treatment (control group). The simulated treatment and control groups are shown in Panel B of Figure IA 3. The simulated control group contains over 4.15 times more borrowers than the simulated treatment group. This is due to two forces: 1) the concentration of borrowers above $200K threshold 2) firm debt below the threshold moves faster and hence firms are more likely to exit the sample range. This simple model makes the point that the discrepancy between the number of firms in the treatment and control sample can be explained if debt growth dynamics are related to the $200K threshold.
Figure IA 1
Distribution of All Borrowers in the Registry Database
This Figure plots the distribution for the entire population of borrowers in March 1998 and March 1999

Panel A: Borrowers with Debt from $100K to $300K in March 1998

Panel B: Borrowers with Debt from $100K to $300K in March 1999
Figure IA 2
Counterfactual Sample in March 1999

Firms are classified into the treatment and control groups using the same criteria as in Figure 1 in the paper, but using 1999 data instead of 1998.
Figure IA 3
Simulated Distribution of Borrowers in March 1998

This plots the simulated distribution of borrowers using the model of debt growth in the Internet Appendix. The simulation is created by starting with a uniform distribution of 200,000 borrowers with debt between $0 and $1,000,000 in January 1997. Each month debt evolves according to the model using the following parameters: $g(I=0)=5\%$, $g(I=1)=1\%$, and $\alpha=1/24$. Panel A plots the simulated population of all borrowers around the $200,000 threshold in March 1998. Panel B plots the simulated treatment and control group in March 1998. A firm is included in the treatment (control) group if its simulated debt between January 1998 and March 1998 is always between $150,000 and $200,000 ($200,000 and $250,000). Both panels show the average of 20 simulations to eliminate noise.

Panel A. Simulated Population of Borrowers

Panel B. Simulated Treatment and Control Group
Using Firm Sample in the $175,000 to $225,000 Total Debt Range

Difference-in-differences (DD) effect of the registry expansion announcement (interim period) and public information (post-expansion period) on (log) debt levels, using specification (5) estimated with OLS:

$$\ln(Debt_{it}) = \alpha + \gamma + \gamma_{Interim} \cdot Interim_{t} + \gamma_{Post} \cdot Post_{t} + \varepsilon_{it}$$

Treatment (control) firms are those with total debt between $175,000 and $200,000 ($200,000 and $225,000) during the pre-announcement period (January to March 1998). Sample: firms whose highest (worst) risk rating during the pre-announcement period is a 2, and with at least one rating of 2 (firms with only good ratings excluded). Columns 1 through 3 are estimated over the subsample of firms with multiple lenders, and column 4 on the subsample with a single lender, before the expansion announcement. Dependent variables: (log) debt of borrower $i$ at time $t$ with the banks that assigned the worst rating (column 1), with the banks that assigned the best rating (column 2), and with all lenders (columns 3 and 4). Right-hand side variable of interest: interaction between a dummy equal to one if borrower $i$ was in the treatment group (information not shared before registry expansion), and dummies equal to one during the interim period and post-expansion period. Coefficients $\gamma_{Interim}$ and $\gamma_{Post}$ are the DD estimates of the effect of the registry expansion announcement and actual expansion on debt. Standard errors are heteroskedasticity robust and clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels.

<table>
<thead>
<tr>
<th>Subsample by # of Lenders before April</th>
<th>Multiple Lenders</th>
<th>Single Lender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ln(Debt from Banks w/ Rating = 2$_{it}$)</td>
<td>ln(Debt from Banks w/ Rating = 1$_{it}$)</td>
</tr>
<tr>
<td>Effect on Debt Level – Interim Period (2 months)</td>
<td>-0.127 (0.212)</td>
<td>-0.069 (0.086)</td>
</tr>
<tr>
<td>Effect on Debt Level – Post Expansion (12 months)</td>
<td>-0.395 (0.298)</td>
<td>-0.216 (0.177)</td>
</tr>
<tr>
<td>Firm Fixed Effects and Month Dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations (Firm-Month)</td>
<td>842</td>
<td>714</td>
</tr>
<tr>
<td>Clusters (Firms)</td>
<td>54</td>
<td>52</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.873</td>
<td>0.778</td>
</tr>
</tbody>
</table>