The Real Costs of Washing Away Corruption:

Evidence from Brazil's Lava Jato Investigation

Claudio Ferraz[a], Luiz Moura[b], Lars Norden[b,c], Ricardo Schechtman[d]

[a] *University of British Columbia, Canada; PUC-Rio de Janeiro, Brazil*

[b] *Brazilian School of Public and Business Administration, Getulio Vargas Foundation, Brazil*

[c] *EPGE Brazilian School of Economics and Finance, Getulio Vargas Foundation, Brazil*

[d] *Central Bank of Brazil, Brazil*

Abstract

Anti-corruption investigations aim at promoting allocative efficiency, growth and innovation, but, if too disruptive, they can generate adverse economic consequences. We examine the costs of one of the world's largest anti-corruption crackdowns, Operação Lava Jato in Brazil, using unique bank-firm-worker data. We find investigated firms cut employment and wages and lose access to bank credit. Importantly, more exposed banks reduce credit also to non-investigated firms, and even more so for politically connected existing borrowers. We further document negative real and financial effects for non-investigated firms more exposed through their banks. Policy makers should consider these costs when devising anti-corruption investigations.

This version: March 14, 2023

Keywords: Corruption, anti-corruption investigations, credit channel, bank lending, capital and labor allocation

JEL classification: D62, D73, G21, G30, O10

1. Introduction

Corruption hampers economic growth, innovation and democracy (Shleifer and Vishny, 1993; Mauro, 1995; Svensson, 2005). However, little is known about the effects of anti-corruption policies, which can have mixed results on the economy as they prosecute and punish corrupt firms. On the one hand, getting rid of corrupt practices allow more efficient and innovative firms to grow and resources to be reallocated to new firms that do not need to be friends with the government. On the other hand, the prosecution of large corrupt firms may disrupt economic activity and generate large adverse economic consequences that may reduce the overall support for future anti-corruption efforts. Most of the literature on the effects of anti-corruption crackdowns focuses on the government sector. In this paper, we focus on the corporate sector and the credit channel. It is well known that credit acts as an amplifier of economic shocks (e.g., Khwaja and Mian, 2008; Bentolila, Jansen, and Jimenez, 2018; Amiti and Weinstein, 2018; Alfaro, García-Santana, Moral-Benito, 2021). We seek to investigate the real costs of one of the largest anti-corruption crackdowns in the world, the *Operação Lava Jato* (Car Wash Operation) in Brazil. We differentiate between the direct effects on the investigated firms and the indirect effects on the rest of the economy.

Brazil provides a unique testing ground to address these questions. The *Operação Lava Jato* aimed at uncovering overbilling and bribery and quickly turned into the largest and most complex anti-corruption investigations in Latin America (Campos et al., 2021). It was a joint operation by the Brazilian Federal Police and the *Ministério Público Federal* (MPF, Federal Prosecution Office of the Public Ministry) and involved 42 billion BRL related to corruption, 6.4 billion BRL bribes directly paid to public officials, over 1 billion documents seized and more than 900 search and seizure warrants in 80 phases during 2014-2021 (e.g., Netto 2016; Campos et al. 2021). Political and economic uncertainty have increased substantially as the

investigations unfolded and the fear of contagion became eminent both in financial markets and in the real economy.

We make two important conjectures in our paper. First, we conjecture that Brazil's large bank based financial system makes bank credit an important natural candidate for the transmission of the Lava Jato scandal to the corporate sector. After the start of the investigations, the affected firms may experience increasing credit constraints and ultimately lose access to bank credit. Second, the rest of the corporate sector may experience indirect effects of Lava Jato. We conjecture that banks that are ex-ante more exposed to Lava Jato firms may grant more or less credit to other firms after the onset of the investigations due to different reasons related to the scandal.

One the one hand, following the anti-corruption campaign, banks might increase lending to "clean" (non-investigated) firms (e.g. Giannetti, Liao and Yu, 2021). Such positive indirect reaction by banks would represent a reallocation of credit, resulting from the surplus generated by reduced or denied credit to corrupt firms. On the other hand, banks that are strongly exposed to investigated firms may reduce credit to non-investigated firms as well because of higher expected losses on credit to Lava Jato firms and/or higher uncertainty about the scale and scope of the investigations. Furthermore, banks' reactions may also depend on concerns about firms from other industries being caught in the scandal. They may cut credit to likely corrupt firms they suspect to become the next targets of investigations. More exposed banks may also suffer from higher risk aversion, increased pressure from market discipline or elevated concerns about their charter values. Note that we do not differentiate between these reactions as they are not mutually exclusive. Our goal in this paper is to provide an estimate of the net effect.

Our analysis is based on a rich dataset that we build from four sources. The first and main source is the Brazilian Credit Information System SCR (*Sistema de Informações de Crédito*) from the Central Bank of Brazil, containing confidential information on virtually all loans made

by financial institutions in Brazil. The second one is RAIS (*Relação Annual de Informações Sociais*) from the Brazilian Ministry of Economics and Labor, which contains information on all formally employed workers. The third one comes from the TSE (*Tribunal Superior Eleitoral*) and contains detailed information on firms' connection to the government through campaign donations in federal elections. The fourth one is the list of firms that have been investigated in the *Operação Lava Jato*, as published by the Public Ministry. We focus on a core set of firms from this list. These firms are large, not publicly listed, from the construction sector and with business relationships to the government. We merge the data from these four sources to obtain a firm-quarter dataset and a more detailed firm-bank-quarter dataset covering the period from 2011 to 2016.

We find the following main results. First, we document that firms that were involved in the Lava-Jato corruption scandal experience significantly negative real effects on firm employment and wage bill. We show that these firms largely lose access to credit: they receive less credit, lower credit ratings and display higher loan loss provisions after the start of the investigations. We established these results in a difference-in-difference analysis, using the start of the anti-corruption investigations as treatment variable. The results are robust to different control groups and different matched samples with similar ex-ante key characteristics and parallel trends in the pre-period. Second, because these firms are large and significant fraction of bank's credit portfolios, we show that banks more exposed to the corruption scandal reduce credit also to non-investigated firms. Hence, there are negative spillover effects of the anti-corruption investigations to the rest of the corporate sector. Third, the reduction of credit by exposed banks is stronger for politically connected existing borrowers. Fourth, for non-investigated firms that borrow from banks with a large ex ante exposure to the shock, we also find negative real and financial effects.

Our paper contributes to the literature in various ways. First, there is a literature on anti-corruption crackdowns, municipality audits and the effects of blacklisting firms. These studies focus on how firms are affected and change their behavior when they lose access to government contracts and the consequences for the local economy and the labor market (Ferraz and Finan, 2008; Ferraz, Finan and Szerman, 2016; Colonnelli, Lagaras, Ponticelli, Prem and Tsoutsoura, 2022; Colonnelli and Prem, 2022; Szerman, 2023).

Another strand of literature focuses on corruption in bank lending. The studies investigate how corruption in bank lending and political connections influence the allocation of credit to firms (Beck, Demirgüç-Kunt and Levine, 2006; Charumilind, Kali and Wiwattanakantang, 2006; Barth, Lin, Lin, and Song, 2009; Weill, 2011; Qi and Ongena, 2019).

Recent research has examined the anti-corruption investigations in China (Chen and Kung, 2019; Giannetti, Liao, You and Yu, 2020; Griffin, Liu and Shu, 2021; Li, Wang and Zhou, 2022). The findings suggest positive effects of the Chinese anti-corruption campaign. However, these results largely disappear when the ruling party or their officials are involved (Griffin, Liu and Shu, 2021). We note it is not easy to extrapolate the evidence from China to other countries because of its political system and a largely state-owned economy.

Our paper differs from the literature in several dimensions. First, the size of the anti-corruption investigations in Brazil as well as the size of the investigated firms are unprecedented. Second, most of the related studies find generally positive effects of anti-corruption investigations on the economy, while we find significantly negative spillover effects on non-investigated firms. Finally, because the Lava Jato investigations represent a significant shock to banks, our paper also relates to the broader literature on the diffusion of economic shocks through banks (e.g., Khwaja and Mian, 2008; Amiti and Weinstein, 2018; Alfaro, García-Santana and Moral-Benito, 2021; Gutierrez, Jaume and Tobal, 2022).

The remainder of this paper is organized as follows. Section 2 provides a brief overview of the institutional characteristics of the *Operação Lava Jato* and presents our main hypotheses. Section 3 describes the data, methodology and summary statistics. Section 4 presents our results on the impact of anti-corruption investigations on credit to investigated firms, credit reallocation and real effects. Section 5 concludes.

2. Institutional background of the Operação Lava Jato

The *Operação Lava Jato*,[1] which started in March 2014 and headed by the Federal Police and the *Ministério Público Federal*, initially investigated money laundry and bribery by a small group of black-market foreign currency dealers that were involved in money laundry, then expanded within a few months to the state-owned oil company Petrobras[2] and the largest Brazilian construction companies that served as its contractors. The operation eventually reached politicians, political parties, state governors, the congress (presidents of both chambers), the federal government of Brazil and even governments of other countries. Essentially, it investigates crimes of active and passive corruption, fraudulent exchange operation, large-scale bribery, kickbacks and an illegal campaign financing scheme of government parties. *Operação Lava Jato* was the largest anti-corruption and anti-money laundry investigations in Brazil and the largest detected corruption scandal in the history of Latin America: it issued more than one thousand warrants for search and seizure, temporary arrest, preventive detection and coercive conduct, aiming at investigating a money laundry scheme that moved billions of *Brazilian Reais* in bribes.

[1] The name *Operação Lava Jato* (Car Wash Operation) is due to a gas station that was used to move illegal values and that was investigated in the first phase of the operation, in which a black-market foreign currency dealer was arrested. Subsequently, the investigations uncovered a direct connection with the former procurement director of Petrobras, who was arrested preventively in the second phase.

[2] Intriguingly, Petrobras was previously seen as "the most autonomous and corporately coherent organization within the Brazilian state enterprise system" (Evans, 1989), an exception if compared to typical glitches of public or state-owned enterprises.

The operation had initially focused on black market foreign currency dealers who used small businesses such as gas stations and car washes to launder money. During the investigations, prosecutors argued that the same criminals laundered money for key executives of Petrobras[3] that were linked to politicians and government parties in an intricate web of corruption. In November 2014, the operation hit a core set of large Brazilian construction companies, including Construtora OAS, Camargo Corrêa and Queiroz Galvão. Shortly afterwards, two further construction companies were added to the list: Andrade Gutierrez and Odebrecht, the latter known as Latin America's largest construction conglomerate (see, for details, Campos et al., 2021).

Essentially, overbilling[4] of contracts for oil refineries, oil rigs, off-shore exploration vessels and office buildings were diverted to secret accounts that shifted the pre-defined percentages of the surplus to politicians, political parties and the corporate conglomerates that were part of the scheme. Billions of U.S. dollars were paid through a web of corruption, in which private interests could acquire political concessions, leading participants to bribe officials in several countries in Latin America and Africa[5], concealing illicit funds in Europe and the United States.

The operation had a successful start and worked efficiently until 2016. At that time, its investigations gradually slowed down as it came closer and closer to politics. In 2019, the *Intercept Brazil Portal* disclosed conversations between the former head judge Sergio Moro and prosecutors that questioned the impartiality of the investigations. Afterwards, Lava Jato

[3] According to the investigations, witnesses testified that the construction companies formed a multi-year cartel to share out contracts and pad prices, perhaps extending beyond petroleum to highway and hydropower contracts. This cartel of the contractors for Petrobras had possibly existed for at least 15 years. Considering only the decade between 2004 and 2014, the companies-maintained contracts with Petrobras, which totaled 59 billion Brazilian *Reais* (see Campos et al 2021).

[4] The construction firms and the public counterparts had formed an agreement that ensured guaranteed business on excessively lucrative terms if they agreed to channel a share of between 1% and 5% of every deal to secret funds (see Campos et al. 2021 and Netto 2016).

[5] 14 countries and some of their heads of state were involved including Argentina, Brazil, Chile, Colombia, the Dominican Republic, Ecuador, Guatemala, Mexico, Panama, Peru, and Venezuela (BBC-Brasil 2017).

lost its luster and the *Procuradoria-Geral da República* (Attorney General's Office) announced the dissolution of its original core at the beginning of 2021. It is now conducted by GAECO (*Grupo de Atuação Especial de Combate ao Crime Organizado*), which is a group part of the *Ministério Público Federal* and the operation came to an end after seven years.

Among the 21 investigated construction firms that we analyze in this paper,[6] two went bankrupt (GDK and Schahin), twelve entered a judicial reorganization process (either during our sample period or afterwards)[7]. Furthermore, there are only seven companies whose CEOs or other key executives were not arrested or wanted by the Federal Police. There were no interested buyers for any of the investigated firms and they were seen as "zombies" or "pariahs", suggesting that financial markets did not expect a bright future for them.

3. Data and methodology

3.1. Data sources

The empirical analyses combine three different data sources. The main source of information comes from the Brazilian Public Credit Register (SCR - Credit Information System), a confidential loan level database owned and managed by the Central Bank of Brazil (BACEN). It contains detailed information on almost all loans in the economy at a monthly level, including loan amounts, interest rates, loan loss provisions, maturities, regulatory borrower ratings and others. Borrower-level characteristics[8], however, are relatively scarce. Therefore, to account for time-invariant or time-varying heterogeneity in firm characteristics, we use firm fixed effects or interacted firm and time fixed effects. Another strength of our firm-

[6] There are 23 construction firms under investigation but only 21 are borrowers and thus included in the credit registry of the Central Bank of Brazil.

[7] The firms that entered in judicial reorganization are: Odebrecht, OAS, Queiroz Galvão, UTC Engenharia, Engevix, IESA, Mendes Junior, Galvão Engenharia, GDK, Schahin, Alumini, and Tomé Engenharia.

[8] Borrower-level information gathered from other sources is limited to firm's number of employees, wage bill, location, age, industry and whether they are publicly listed or not.

bank-time data is that we are able to distinguish the existing borrowers (intensive margin) from the new borrowers (extensive margin) and investigate whether banks behaved differently in their response to the anti-corruption investigations across these two groups.

We also use data from *Relação Anual de Informações Sociais* (RAIS), an administrative data set collected on an annual basis by the Brazilian Ministry of Economics and Labor, which covers all formal workers in Brazil. The sample we use is restricted to large firms having more than 250 employees in 2012.

Our third dataset provided by the *Tribunal Superior Eleitoral* (TSE) refers to campaign contributions in the federal elections of 2010, with detailed information about donors' contributions and recipients. For each candidate, beyond the identification of the parts involved in the contribution, it is possible to identify the political party, the state, position of the candidate (state deputy, federal deputy, senator, governor or president) and the size of the campaign contribution in the election years.

The fourth dataset comes from IBGE (*Instituto Brasileiro de Geografia Estatística*) and has detailed macroeconomic indicators at the yearly state-level from 2012 to 2016 such as sales volume, nominal revenue, industrial production, unemployment rates and human development index. We also gathered sectoral indices of innovation both from PINTEC (IBGE Innovation Survey) and from OECD (Organisation for Economic Co-operation and Development). These characteristics were matched to each firm according to their headquarters' location or industry.

The merged dataset comprises free-market credit[9] granted in the period from January 2012 to July 2016[10] of large firms (with more than 250 employees in 2012). State-owned banks are excluded from the initial analysis because they might have counter-cyclical behavior in periods

[9] Free market lending, in contrast to earmarked lending, refers to the type of credit that does not meet any public sector directions or has subsidized interest rates.

[10] Our sample period ends before the second semester of 2016 before the impeachment of the president Dilma Rousseff, which introduced substantial economic and political uncertainty in the country.

of credit shrinkage in the economy (Capeleti, Garcia, Miessi, 2022). Financial firms and state-owned firms are also excluded.

3.2 Methodology

We first conduct a difference-in-difference (DID) analysis to study the direct effects of the anti-corruption investigations on Lava Jato firms. We split the sample into two periods: one before the start of the investigations (2013) and one afterwards (2014, 2015, and the first 2 quarters of 2016).[11] We employ aggregate data at the firm-quarter level. To test the effects of Lava Jato on credit availability and credit terms for the investigated firms, we estimate the difference-in-differences model shown in equation (1):

$$C_{i,t} = \alpha + \beta_1 Lava\ Jato_i + \beta_2 Post_t + \beta_3 (Lava\ Jato_i \times Post_t) + \upsilon_i + \theta_t + \varepsilon_{i,t} \quad (1)$$

where $C_{i,t}$ stands for real or financial outcomes of firm i at time t such as wage bill, number of employees, firm growth, as well as information on new loans, borrower rating, loan loss provision, maturity and pre-fixed interest rate. $Lava\ Jato_i$ indicates each of the 21 investigated construction firms. $Post_t$ is a dummy variable that indicates the period after the start of the anti-corruption investigations. We exclude the first quarter of 2014 because the anti-corruption investigations started in the middle of March 2014.[12] We control for time-invariant firm characteristics by including firm fixed effects υ_i. We also control common macroeconomic shocks to firms that may change over time using time fixed effects θ_t and the standard errors are clustered at the firm level.

[11] In unreported robustness tests, we employ a symmetric time period that considers two years before Lava Jato (2012-2013) and two years after Lava Jato (2014-2015). The results are qualitatively similar.

[12] We consider this time period (2013Q1-2016Q2, except 2014Q1) in all regression analyses of this paper.

We then investigate potential spillover and reallocation in the credit market. Banks with credit portfolios more exposed to investigated firms may grant more or less credit to non-investigated borrowers after the onset of the investigations. The effects depend on the surplus from denied or reduced credit to investigated firms, expected losses on outstanding credit to investigated firms and expectations about the dynamics of the anti-corruption investigations. We estimate the indirect effect of the anti-corruption investigations on non-investigated firms using the model (2). We estimate this model at the bank and bank-firm level.

$$C_{i,j,t} = \alpha + \beta_1 \textit{Lava Jato bank exposure}_j + \beta_2 \textit{Post}_t + \beta_3(\textit{Lava Jato bank exposure}_j \times \textit{Post}_t) + X_{j,t-1} + \upsilon_{i,t} + \varphi_j + \varepsilon_{i,j,t} \quad (2)$$

Where $C_{i,j,t}$ is either the amount of new loans for firm *i* from bank *j* in quarter *t*, at the intensive margin, or a dummy indicating whether firm *i* received a new loan with bank *j* in quarter *t*, at the extensive margin. *Lava Jato bank exposure*$_j$ is a dummy that equals one if bank *j* has a high share (upper tercile) of outstanding credit to firms in 2012 that become subsequently investigated in the *Operação Lava Jato* (see equation 3).

$$\textit{Lava Jato bank exposure}_j = \frac{\sum \textit{Outstanding credit}_i}{\sum \textit{Outstanding credit}_{ni} - \sum \textit{Outstanding credit}_i} \quad (3)$$

Note that the computation of a bank's *Lava Jato bank exposure* is based on data from 2012, while the estimation period ranges from 2013 to mid-2016. Importantly, we consider data on investigated firms to create the bank exposure measure but estimate the regression only on data from non-investigated firms. The subscript *j* denotes banks, *i* stands for investigated firms, and *ni* denotes non-investigated firms. *Post*$_t$ refers to the period after the investigations. The key term in our analysis is the interaction term *Lava Jato bank exposure* × *Post*. The coefficient β_3

indicates the DID estimator. We include lagged bank characteristics $X_{j,t-1}$ as control variables and a set of either firm-time $\upsilon_{i,t}$ and bank fixed effects φ_j or firm-bank and time fixed effects (not shown in equation 1). Firm-time fixed effects purge all time variation in the data that at the firm-level and captures any determinants of firm credit demand, allowing us to isolate supply factors. Firm-bank fixed effects control for unobserved bank-firm relationship characteristics. We cluster the standard errors at the bank-time level. Moreover, next to the likelihood of a new loan or the amount of new loans, we investigate other characteristics of loans to non-investigated firms such as credit rating, interest rate, maturity and loan loss provisions.

We then extend the previous specification by interacting bank exposure with characteristics of the borrowing firms to test the heterogeneity of credit reallocations across firms. We estimate the regression model shown in equation (4):

$$\begin{aligned} C_{i,j,t} = \alpha + \beta_1 &\textit{Lava Jato bank exposure}_j + \beta_2 \textit{Post}_t + \beta_3 \textit{Moderator}_i \\ &+ \beta_4(\textit{Lava Jato bank exposure}_j \times \textit{Post}_t) + \beta_5(\textit{Moderator}_i \times \textit{Post}_t) \\ &+ \beta_6(\textit{Lava Jato bank exposure}_j \times \textit{Moderator}_i \times \textit{Post}_t) + X_{j,t-1} + \upsilon_{i,t} + \varphi_j + \varepsilon_{i,j,t} \quad (4) \end{aligned}$$

All variables are similar to the previous model, except the firm-level moderator. This variable considers ex-ante borrower characteristics such as the size of donations in the federal elections of 2010 (which we employ as a proxy for the connection with the government) a dummy indicating firms from the construction sector, and other firm-characteristics fixed in 2012, such as number of employees, mean wage, age, and number of bank-relationships measured by the end of 2012. The same approach is used for macroeconomic characteristics. This model also includes a set of lagged bank controls X, and firm-time and bank fixed effects, we cluster the standard errors at the bank-time level. In an unreported analysis, we employ

firm-bank and time FE and obtain similar results. The same model allows to investigate potential geographic credit reallocation in Brazil by including state-level characteristics as moderators (also fixed in the year 2012), such as state sales volume, nominal revenue, industrial production, and unemployment rate.

Finally, we examine whether non-investigated firms that are indirectly, i.e., through their bank relationships in the pre-period, more exposed to Lava Jato firms suffer any real effects. We consider key labor market outcomes *L* such as employment, wage bill and mean wage paid. We estimate the regression model shown in equation (5):

$$L_{i,t} = \alpha + \beta_1 Lava\ Jato\ firm\ exposure_i + \beta_2 Post_t + \beta_3 (Lava\ Jato\ firm\ exposure_i \times Post_t) + \varepsilon_{i,t} \quad (5)$$

Where $L_{i,t}$ stands for wage bill, number of employees, and mean wage from firm *i* in year *t*. *Lava Jato firm exposure*$_i$ indicates borrower *i's* exposure to Lava Jato firms through its bank relationships in the previous period.[13] The variable, measured in 2012, is defined as follows:

$$Lava\ Jato\ firm\ exposure_i = \frac{\sum (Lava\ Jato\ Exposure_j \times Outstanding\ credit_{i,j,t})}{\sum Outstanding\ credit_{i,t}} \quad (6)$$

Model (5) includes lagged firm controls. Standard errors are clustered at the firm level.

3.3. Summary statistics

Table 1 reports summary statistics of the main variables used in this paper.

[13] Similar to the bank-level exposure to Lava Jato firms, we measure the *Firm Exposure* prior to the start of our sample period to ensure its exogeneity.

(Insert Table 1 here)

For the analysis of investigated firms, *New Loans*$_{it}$ indicates that, on average, firms borrow R$20,541,316 per quarter in new loans. Alternatively, for the analysis including non-investigated firms, *New Loans*$_{ijt}$ and *New Loans [dummy]*$_{ijt}$ indicates that existing borrowers borrow, on average, R$1,734,128 per bank-quarter in new loans and about 2.4% of the new contracts are new loans, respectively. As expected, the median value for both the continuous and the dichotomous variables are zero, since we mechanically included zeros for absent new loans in case there exists a stock of credit for that firm-bank notch (intensive margin) or included zeros for all missing new loans (extensive margin). Additionally, the loans are usually paid back in around 11 months (median: 5), rating has a moderate grade of 2, the pre-fixed interest rate is much bigger than the mean cost of capital during the same period (mean: 47.05%, median: 19.64%), and banks usually set aside 0.89% of their portfolio as provisions to account for future losses on loan defaults (median: 0.49%).

Moreover, non-investigated firms are large (mean: 1,035 employees; median: 485 employees) and with relatively skilled workers with wages (mean: R$1,992.72/month, median: R$1,546.53/month) well above the minimum wage in the country[14]. The firms donate an average of R$62.67 (median: R$0) per employee in the federal elections of 2010 and are well connected in the banking system (mean: 3.52 bank relationships in 2012).

4. Results

4.1. Impact of Lava Jato on investigated firms

We start our analysis by examining the effects of Lava Jato on investigated firms. Table 2 presents the results on real effects.

[14] The minimum wage in Brazil during our sample period was R$724 in January 2014 and R$880 in July 2016.

(Insert Table 2 here)

We find highly significant real negative effects of the anti-corruption investigations. Because the coefficients of interest are high and the related independent variables are discrete, the semi-elasticity is better approximated exponentiation of the coefficient and subtracting by one. In that sense, the wage bill decreases by 67% and number of employees decreases by 61% for investigated firms after the onset of the investigations. The adverse effects increase monotonically over time as we see in columns 3 and 4 or alternatively in the graphs of the coefficients with confidence intervals at Figure 1 where 2011 is defined as the baseline year and estimations include a different post dummy for each year afterwards. Figure 1 shows that before 2014 there is no differentiation between investigated and non-investigated firms, consistent with the assumption of parallel trends.

(Insert Figure 1 here)

Next, we study the financial effects starting in Table 3 Panel A with the changes in the amount of new loans. Column 1 represents the baseline model whereas specifications 2 and 3 make use of alternative matching sample strategies for the control groups (see Appendix, Table A1).

(Insert Table 3 here)

The coefficients of the interaction terms are quite high and represent almost a complete depletion of new loans if one makes the exponentiation. However, in fact the estimates

magnitudes are not very good summaries because some Lava Jato firms do experience nearly a complete depletion while others not (and besides the transformation log (1+.) complicates the interpretation further) but the overall message is that the effects are huge.

Figure 2 Panel A shows the interaction coefficients of the previous models split by quarter, using the 2012Q1 as the reference period.

(Insert Figure 2 here)

We see that there are no differential trends between the two groups of firms before the onset of the investigations. And although the interaction coefficients start a decreasing pattern after that, they only become significant a year later in 2015 and intensify overtime until 2016Q1, consistent with the fact that Lava Jato firms were mostly hit by the investigations by the end of 2014.

As an additional robustness exercise, we also perform a synthetic control estimation first aggregating new loans by industry and then proceeding to create a synthetic control based on other industries apart from construction. Figure 2 Panel B shows that Lava Jato firms and the synthetic control behave quite similarly before the investigations but there is a sizable contraction of credit to LJ firms after 2014, reaching an almost depletion in 2016 Q1

Finally, in Table 3 Panel B we take loan terms as the dependent variables. We find that regulatory ratings given by banks to LJ firms deteriorate by almost one category (larger ratings are worse), and similarly loan loss provisions increase 0.6%, which is material compared to the average of 0.9% in the whole sample. Maturities also decrease and interest rates increase but they are not statistically significant.

4.2. The effects of Lava Jato on non-investigated firms at the bank level

In the next step, we investigate the indirect effects of Lava Jato on non-investigated firms, performing an analysis at the bank level.

Figure 3 shows the *Lava Jato bank exposure*, which is the credit exposure of privately owned banks to Lava Jato firms (blue bars) and the number of firm-quarter observations of each bank (line), both measured in 2012.

(Insert Figure 3 here)

Banks are ranked by the size of their exposure to Lava Jato firms. There is substantial heterogeneity across individual banks. The three banks with the largest exposures display such values between 8% and 15% of their portfolios. Moreover, the biggest banks in Brazil, as measured by the peaks in the number of firm-quarter observations in our sample, have just a moderate level of LJ exposures. The upper tercile of the *Lava Jato bank exposure* distribution contains the banks to the left of the dotted blue line. There is one large bank and some medium banks there.

(Insert Table 4 here)

Table 4 shows that banks with greater ex-ante exposure to Lava Jato firms decrease lending to non-investigated firms more than other banks after the onset of the investigations and the results are robust regardless of the exposure measure adopted (discrete or continuous) and are concentrated in the upper tercile instead of the middle tercile.

(Insert Table 5 here)

Considering the decrease of credit to the corporate sector it is possible that banks would increase lending to the household sector. Using another dataset that aggregates lending by borrower type, we show in Table 5 that this is not the case. The only statistically significant interactions show up when the response variable is (log of) new loans granted to the business sector, columns (4) to (6).

4.3 The effects of Lava Jato on non-investigated firms at the firm-bank level

We now investigate the indirect effects of Lava Jato on non-investigated firms, performing a more granular analysis at the firm-bank level.

(Insert Table 6 here)

Table 6 shows that new borrowers have a lower chance of getting new loans from more exposed banks, regardless of the fixed effects added to the models (firm-time and bank; or firm-bank and time). The decrease in probabilities might seem small but they are not, since they refer to all potentially new borrowers in our sample, not necessarily to those who have applied for new loan (we don't have information on loan applications). Current borrowers also receive less credit from more exposed banks after the onset of the investigations (intensive margin) but those results are not very significant, so that on average our results appear to be driven mostly by the extensive margin.

In an additional analysis, we examine whether the timing of the Lava Jato exposure influences our results. If we measure the exposure too close to the start of the anti-corruption investigations, it might be affected by anticipation effects by banks. If we measure it too much back in the past, it might be outdated and uninformative. We therefore measure the Lava Jato bank exposure in 2011 (instead of 2012) and re-estimate the same regression models. The

corresponding results, shown in the Appendix, Table A2 are quantitatively even stronger and highly significantly negative for the extensive and intensive margin.

We now dig deeper to provide more evidence on banks' reaction to the Lava Jato shock. In Table 7, we interact the main variables with (the third tercile dummy of) a measure of election campaign donations by each firm, a proxy for government connection and perhaps also related to undetected corruption. In fact, in a time when the whole country was expecting who was the next firm going to be caught in the Lava Jato scandal, suspicion could have rested more in firms that had donated a lot to future government representatives.

(Insert Table 7 here)

As before, we see that the double interaction *Lava Jato bank exposure x Post* is significant only for the extensive margin. Additionally, firms that have donated more have, after the start of the investigations, a lower chance of getting new credit and receive less credit when they are already current borrowers (coefficient on *Post x Donations* in columns 2 and 5 respectively). More exposed banks give more new credit to current borrowers that are government connected in general (coefficient on *Lava Jato bank exposure x Donations* in columns 4 and 6) but decrease such new loans to them after the onset of the investigations (triple interaction for the intensive margin). On the other hand, at the extensive margin the triple interaction is not robust and displays a surprising highly significant positive signal in one specification.

(Insert Table 8 here)

In Table 8, we investigate the interaction of the main variables with a dummy variable indicating whether a firm's number of bank relationships is greater than one in 2012. Most of the firms have more than one bank relationship, so it is easier to interpret the relative effects from the point of view of the firms that have only one bank relationship in 2012 (whose effects correspond to the opposite of the signs displayed). Those firms have a higher chance of getting credit as new borrowers from an exposed bank (triple interaction at the extensive margin). However, existing borrowers with only one exposed bank experience reduced lending (triple interaction at the intensive margin). The exclusive bank possibly has more bargaining power as the firm possibly does not have important alternative financing options. Table 8 still shows that exclusive borrowers are generally more linked to more exposed banks during the whole sample and also experience a lower chance of getting new credit from all banks after the onset of the investigations.

We saw earlier that non-investigated firms experienced a decline in new credit from more exposed banks. However, it is also possible that less exposed banks substitute away from this reduced credit, so that the total impact at the firm level would be muted. Therefore, next we investigate indirect effects of Lava Jato on non-investigative firms at the firm level, considering aggregate credit received from each of those firms from the set of all private banks in our sample. We assume that the previously defined *Lava Jato firm exposure* contains the key transmission channel for such analysis.

(Insert Table 9 here)

Table 9 shows that there is also a decrease in new credit at the firm level (at least from the total of private banks in our sample). In column 2, when only positive new loans are considered, the decrease is in the magnitude of 23%. The average regulatory rating and the sum of loan

loss provisions of more indirectly exposed firms also increase and even the average interest rate increases slightly and significantly in column 6. The effects are smaller than the direct effects of Table 3, as expected, but the reduction in credit is still material.

(Insert Table 10 here)

Table 10 estimates the real effects that are possibly related to the unexpected credit crunch associated with the Lava Jato scandal. Firms indirectly more exposed to the scandal through the banking system reduce their wage bill by 11% and their number of employees by 8.5% after the onset of the investigations. Although such figures are far in magnitude from the direct effects of Table 2, of the order of 60%, they are still economically meaningful and, as the former, they also increase over time.

4.4. Further analyses

In this section, we briefly summarize findings from further analyses and additional empirical checks.

First, we provide more evidence on the spillover effects of Lava Jato. We examine whether banks' response to the anti-corruption investigations is related to the firm innovation. This analysis is motivated by prior research showing that corruption hampers investment in innovation and economic development. If corruption is partially mitigated after the beginning of anti-corruption investigations, we expect that more exposed banks should reallocate credit to more innovative firms because of their growth opportunities. We investigate how the reallocation of credit after the start of Lava Jato varies across firms' level of innovation using data provided by the IBGE at the industry level.[15] The baseline category for comparison in our

[15] In unreported analyses, we employ the metric of innovation proposed by OECD and the results remain robust.

specification are firms from industries with low levels of innovation (such as manufacturing of wood products, textiles, beverages, tobacco, leather goods, among others). The Appendix, Table A3 shows the results. We find that banks that are more exposed to Lava Jato reallocate credit significantly more to borrowers from more innovative industries at the intensive margin. The result is stronger for existing borrowers in highly innovative sectors, such as scientific research and development, manufacturing of computer equipment, electronic, optical, chemical, and pharmaceutical products. This result is in line with the literature documenting that corruption undermines innovation (Teece, 1981; Shleifer and Vishny, 1993; Rose-Ackerman, 2001 and 2004). At the extensive margin, nonetheless, firms in industries classified as medium innovation (manufacture of petroleum and biofuel products, rubber and plastic products or printing and reproduction of recordings) or medium-high innovation (information technology services, manufacture of food products, furniture, machinery and equipment, pulp and paper, or electricity, gas and other utilities) are less likely to receive credit from more exposed banks. The findings on the intensive and extensive margin together suggest that banks reallocate credit to existing borrowers that are highly innovative (and not to new borrowers), likely because of validated proprietary information about their innovation strategy, business model or products.

Second, our previous analysis is based on data from privately owned banks in Brazil. We expand the sample and add state-owned banks to the analysis. These banks are important in Brazil as they exhibit a market share of about 40%. They are responsible for state-led lending programs related to economic development, are subject to government influence and exhibit a weaker governance. The Appendix, Table A4 shows the results. Similar to the baseline findings, we find that new borrowers receive on average less credit from more exposed banks in the post-period. However, overall, state-owned banks response to the anti-corruption investigations is less clear than the one of privately owned banks.

5. Conclusions

In this paper, we investigate the effects of one of the world's largest anti-corruption investigations: the *Operação Lava Jato* in Brazil, using unique bank-firm-worker data. We conduct a difference-in-differences analysis of the real and financial effects on investigated and non-investigated firms, considering the credit channel as transmission mechanism.

We find that anti-corruption investigations "work", i.e., they have negative real and financial effects on likely corrupt firms. However, we also find significant negative spillovers of anti-corruption investigations on the rest of the corporate sector (but not on the household sector). This negative effect is stronger for politically connected existing borrowers. We further show negative real and financial effects for non-investigated firms more exposed through their bank relationships.

Our paper has several important implications. Governments should ex ante consider indirect real and financial effects of anti-corruption investigations. We document negative credit spillover effects and negative real effects on labor market outcomes for non-investigated firms in Brazil. These findings suggest that the economic impact of anti-corruption investigations is clearly not as straightforward as implied by evidence from related studies, especially those about China. Bank supervisors and regulators should be aware of the direct and indirect effects and their impact on financial stability. Finally, our findings indicate how spillovers effects in the credit market can result in real effects. Firms should be aware of these effects and take measures to shield themselves against these spillover effects (e.g., high transparency, close bank-firm relationships, loan commitments, etc.). We acknowledge that our results on bank credit reallocation likely underestimate the full effect. There might be further negative effects through trade credit chains in the corporate sector due to direct effects of official debarments (blacklisting) and credit risk spillover.

Appendix

Table A1: Donor industries weights in Synthetic Lava Jato Control Group

Industry code (CNAE)	Industry	Weight
20	Manufacturing of chemicals	.077
29	Manufacturing of motor vehicles, trailers and bodywork	.087
31	Furniture manufacturing	.146
45	Trade and repair of automotive vehicles and motorcycles	.254
91	Activities related to cultural and environmental heritage	.364
93	Sports, recreation and leisure activities	.072
Sum		1.000

Table A2: Effects on credit to non-investigated firms using the LJ Exposure from 2011

This table shows firm-bank level regression results of the model $C_{i,j,t} = \alpha + \beta_1 Lava\ Jato\ bank\ exposure_j + \beta_2 Post_t + \beta_3(Lava\ Jato\ bank\ exposure_j \times Post_t) + \varepsilon_{i,j,t}$ where $C_{i,j,t}$ denotes either New Loans [dummy] or Ln(1+New Loans). The analysis of the extensive margin considers only firms that did not borrow before 2014Q1, while the intensive margin considers only firms that did borrow before 2014Q1. Both analyses exclude Lava Jato firms. *Lava Jato bank exposure*, as defined in Table 1 but here measured in 2011, is a dummy that equals one if the bank is in the upper tercile of the distribution of the continuous exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank-level. Standard errors (in parentheses) are clustered at the bank-time level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)
	Extensive margin		Intensive margin	
Dep. Var.:	*Dummy [New Loans>0]*	*Dummy [New Loans>0]*	*Log(1+New Loans)*	*Log(1+New Loans)*
Lava Jato bank exposure$_{T3}$ × Post	-0.013***	-0.019***	-0.428**	-0.989***
	(0.002)	(0.003)	(0.178)	(0.205)
Bank controls $_{t-1}$	Yes	Yes	Yes	Yes
Firm-Time-FE, Bank-FE	Yes	No	Yes	No
Firm-Bank-FE, Time-FE	No	Yes	No	Yes
Number of observations	203,370	273,728	269,566	324,216
Adjusted-R^2	0.080	0.130	0.153	0.282

Table A3: Effects on credit to non-investigated firms and innovation

This table shows the regression results of the model $C_{i,j,t} = \alpha + \ldots + \beta_4 Lava\ Jato\ bank\ exposure_j \times Post_t + \beta_5 Lava\ Jato\ bank\ exposure_j \times Innovation_i + \beta_6 Post_t \times Innovation_i + \beta_7 (Lava\ Jato\ bank\ exposure_j \times Post_t \times Innovation_i) + \varepsilon_{i,j,t}$ where *C* indicates either the volume of new loans *Ln(1+New loans)* or whether the firm took a new loan or not *New Loans [dummy]*. The analysis of the intensive margin considers only firms that did borrow before 2014Q1 and excludes Lava Jato firms. The analysis of the extensive margin considers only firms that did not borrow before 2014Q1 and excludes Lava Jato firms. *Lava Jato bank exposure* is a dummy that equals one if the bank is in the upper tercile of the distribution of the continuous exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Firms' innovation is a categorical variable proposed by PINTEC that classifies industries according to two types of innovation: product innovation and process innovation. The baseline category for firms' level of innovation is *low innovation*. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank-time level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)
	Intensive margin		Extensive margin	
Dep. Var.:	*Ln(1+New Loans)*	*Ln(1+New Loans)*	*New Loans [dummy]*	*New Loans [dummy]*
Lava Jato bank exposure × Post x Medium-low innovation	0.433	0.936**	-0.00553	0.00997
	(0.355)	(0.371)	(0.00786)	(0.00693)
Lava Jato bank exposure x Post x Medium innovation	0.203	0.525*	-0.0179**	-0.0193***
	(0.349)	(0.298)	(0.00726)	(0.00583)
Lava Jato bank exposure x Post x Medium-high innovation	0.390	0.560**	-0.0209***	-0.0156***
	(0.278)	(0.268)	(0.00432)	(0.00431)
Lava Jato bank exposure x Post x High innovation	0.585	1.183***	-0.0146*	-0.00454
	(0.443)	(0.400)	(0.00755)	(0.00663)
Lava Jato bank exposure x Medium-low innovation	-0.406		0.00231**	
	(0.279)		(0.00101)	
Lava Jato bank exposure x Medium innovation	0.304		-0.00127	
	(0.267)		(0.000835)	
Lava Jato bank exposure x Medium-high innovation	-0.105		5.80e-05	
	(0.221)		(0.000548)	
Lava Jato bank exposure x High innovation	-0.665*		0.00126	
	(0.366)		(0.00103)	
Post x Medium-low innovation		-0.614***		0.00286
		(0.190)		(0.00331)
Post x Medium innovation		-0.163		0.0119***
		(0.190)		(0.00352)
Post x Medium-high innovation		-0.567***		0.00825**
		(0.154)		(0.00239)
Post x High innovation		-0.435*		0.00239
		(0.252)		(0.00283)
Lava Jato bank exposure x Post	-0.316	-0.964***	0.000947	-0.00379
	(0.251)	(0.287)	(0.00417)	(0.00477)
Bank controls [t-1]	Yes	Yes	Yes	Yes
Firm-Time-FE, Bank-FE	Yes	No	Yes	No
Firm-Bank-FE, Time-FE	No	Yes	No	Yes
Number of observations	77,596	90,016	61,876	79,937
Adjusted-R^2	0.170	0.313	0.093	0.164

Table A4: Effects on credit to investigated firms including state-owned banks

This table shows the regression results of the model $C_{i,j,t} = \alpha + \ldots + \beta_4(Lava\ Jato\ Exposure_j \times Post_t) + \beta_5(Lava\ Jato\ Exposure_j \times State\ Owned\ Banks_j) + \beta_6(Post_j \times State\ Owned\ Banks_j) + \beta_7(Lava\ Jato\ Exposure_j \times Post_t \times State\ Owned\ Banks_j) + \varepsilon_{i,j,t}$ where $C_{i,j,t}$ denotes either New Loans [dummy] or Ln(1+New Loans). The analysis of the extensive margin considers only firms that did not borrow (free market lending) before 2014Q1, while the intensive margin considers only firms that did borrow (free market lending) before 2014Q1. Both analyses exclude Lava Jato firms. *Lava Jato exposure*, as defined in Table 1, is a dummy that equals one if the bank is in the upper tercile of the distribution of the continuous exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank-level. Standard errors (in parentheses) are clustered at the bank-time level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)
	Extensive margin		Intensive margin	
Dep. Var.:	*New Loans [dummy]*	*New Loans [dummy]*	*ln(1+New Loans)*	*ln(1+New Loans)*
Lava Jato bank exposure × Post	-0.0102***	-0.0113***	-0.0530	-0.3644*
	(0.0026)	(0.0032)	(0.1569)	(0.1856)
Post x State owned banks	0.0005	-0.0021	-1.2022***	-1.1197***
	(0.0028)	(0.0033)	(0.1797)	(0.2207)
Lava Jato bank exposure × Post × State owned Banks	0.0087	0.0231***	1.5734***	1.8686***
	(0.0061)	(0.0064)	(0.3277)	(0.3670)
Bank controls $_{t-1}$	Yes	Yes	Yes	Yes
Firm-Time-FE, Bank-FE	Yes	No	Yes	No
Firm-Bank-FE, Time-FE	No	Yes	No	Yes
Number of observations	208,826	263,344	269,567	306,870
Adjusted-R^2	0.0774	0.1525	0.1712	0.3222

References

Alfaro, L., García-Santana, M., Moral-Benito, E., 2021. On the Direct and Indirect Real Effects of Credit Supply Shocks. *Journal of Financial Economics* 139, 895-921.

Amiti, M., Weinstein, D., 2018. How Much Do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data. *Journal of Political Economy* 126, 525-587.

Avis, E., Ferraz, C., Finan, F., Varjão, C., 2019. Money and politics: The effects of campaign spending limits on political competition and incumbency advantage. Working Paper, https://www.nber.org/papers/w23508.

Barth, J., Lin, C., Lin, P., Song, F., 2009. Corruption in bank lending to firms: Cross-country micro evidence on the beneficial role of competition and information sharing. *Journal of Financial Economics* 91, 361-388.

Beck, T., Demirgüç-Kunt, A., Levine, R., 2006. Bank supervision and corruption in lending. *Journal of Monetary Economics* 53, 2131-2163.

Beck, T., Demirgüç-Kunt, A., Maksimovic, V., 2005. Financial and legal constraints to firm growth: Does firm size matter? *Journal of Finance* 60, 137-177.

Becker, G., 1968. Crime and punishment: An economic approach. *Journal of Political Economy* 76, 169-217.

Bentolila, S., Jansen, M., Jiménez, G., 2018. When Credit Dries Up: Job Losses in the Great Recession. *Journal of the European Economic Association* 16, 650-695.

Boas, T., Hidalgo, F., Richardson, N., 2014. The spoils of victory: campaign donations and government contracts in Brazil. *Journal of Politics* 76, 415-429.

Campos, N., Engel, E., Fischer, R., Galetovic, A., 2021. The ways of corruption in infrastructure: lessons from the Odebrecht case. *Journal of Economic Perspectives* 35, 171-190.

Capeleti, C., Garcia, M., Miessi, F., 2022. Countercyclical credit policies and banking concentration: Evidence from Brazil. *Journal of Banking and Finance* 143, 106589.

Charumilind, C., Kali, R., Wiwattanakantang, Y., 2006. Connected lending: Thailand before the financial crisis. *Journal of Business* 79, 181-218.

Chen, T., Kung, J., 2019. Busting The 'Princelings': The Campaign Against Corruption in China's Primary Land Market. *Quarterly Journal of Economics* 134, 185-226.

Claessens, S., Feijen, E., Laeven, L., 2008. Political connections and preferential access to finance: The role of campaign contributions. *Journal of Financial Economics* 88, 554-580.

Colonnelli, E., Lagaras, S., Ponticelli, J., Prem, M., Tsoutsoura, M., 2022. Revealing corruption: Firm and worker level evidence from Brazil. *Journal of Financial Economics* 143, 1097-1119.

Colonnelli, E., Prem, M., 2022. Corruption and firms. *Review of Economic Studies* 89, 695-732.

Evans, P., 1989. Predatory, developmental, and other apparatuses: A comparative political economy perspective on the third world state. *Sociological Forum* 4, Kluwer Academic Publishers-Plenum Publishers.

Ferraz, C., Finan, F., 2008. Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes. *Quarterly Journal of Economics* 123, 703-745.

Ferraz, C., Finan, F., Szerman, D., 2016. Procuring firm growth: the effects of government purchases on firm dynamics. NBER Working Paper 21219.

Fisman, R., Guriev, S., Ioramashvili, C., Plekhanov, A., 2021. Corruption and firm growth: evidence from around the world. Working paper, SSRN 3828225.

Giannetti, M., Liao, G., You, J., Yu, X., 2020. The Externalities of Corruption: Evidence from Entrepreneurial Activity in China. *Review of Finance* 25, 629-667.

Griffin, J., Liu, C., Shu, T., 2021. Is the Chinese Corporate Anti-Corruption Campaign Authentic? *Management Science*, https://doi.org/10.1287/mnsc.2021.4181.

Khwaja, A., Mian, A., 2008. Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market. *American Economic Review* 98(4), 1413-1442.

Mauro, P., 1995. Corruption and growth. Q*uarterly Journal of Economics* 110, 681-712.

Netto, V., 2016. Lava Jato: o juiz Sergio Moro e os bastidores da operação que abalou o Brasil. Rio de Janeiro: *Primeira Pessoa*.

Qi, S., Ongena, S., 2019. Will money talk? Firm bribery and credit access. *Financial Management* 48, 117-157.

Rose-Ackerman, S., 2001. Trust, honesty, and corruption: reflection of the state-building process. *European Journal of Sociology* 42, 27–71.

Rose-Ackerman, S., 2004. The challenge of poor governance and corruption. *Revista Direito GV*, 1.5 Especial 1.

Shleifer, A., Vishny, R., 1993. Corruption. Q*uarterly Journal of Economics* 108, 599-617.

Svensson, J., 2005. Eight questions about corruption. *Journal of Economic Perspectives* 19, 19-42.

Szerman, C., 2023. The Employee Costs of Corporate Blacklisting: Evidence from Brazil. *American Economic Journal: Applied Economics* 15, 411-41.

Teece, D., 1981. The market for know-how and the efficient international transfer of technology. *Annals of the Academy of Political and Social Science*, 81–96.

Weill, L., 2011. How corruption affects bank lending in Russia. *Economic Systems* 35, 230-243.

Figure 1: Real effects on Lava Jato firms

The figures display regression coefficients of the model $C_{i,t} = \alpha + \beta_1 Lava\ Jato_i + \beta_2 Post_t + \beta_3 (Lava\ Jato_i \times Post_t) + \varepsilon_{i,t}$ where *C* indicates either *Wage bill or Number of employees. Lava Jato* is a dummy variable that equals one for the 21 construction firms that are subject to anti-corruption investigations and zero otherwise. *Post* is represented by quarter dummies indicating the effect over time. Graphs in Panel A consider 2011 as baseline category, while graphs in Panel B consider 2012 as baseline category.

Panel A **Panel B**

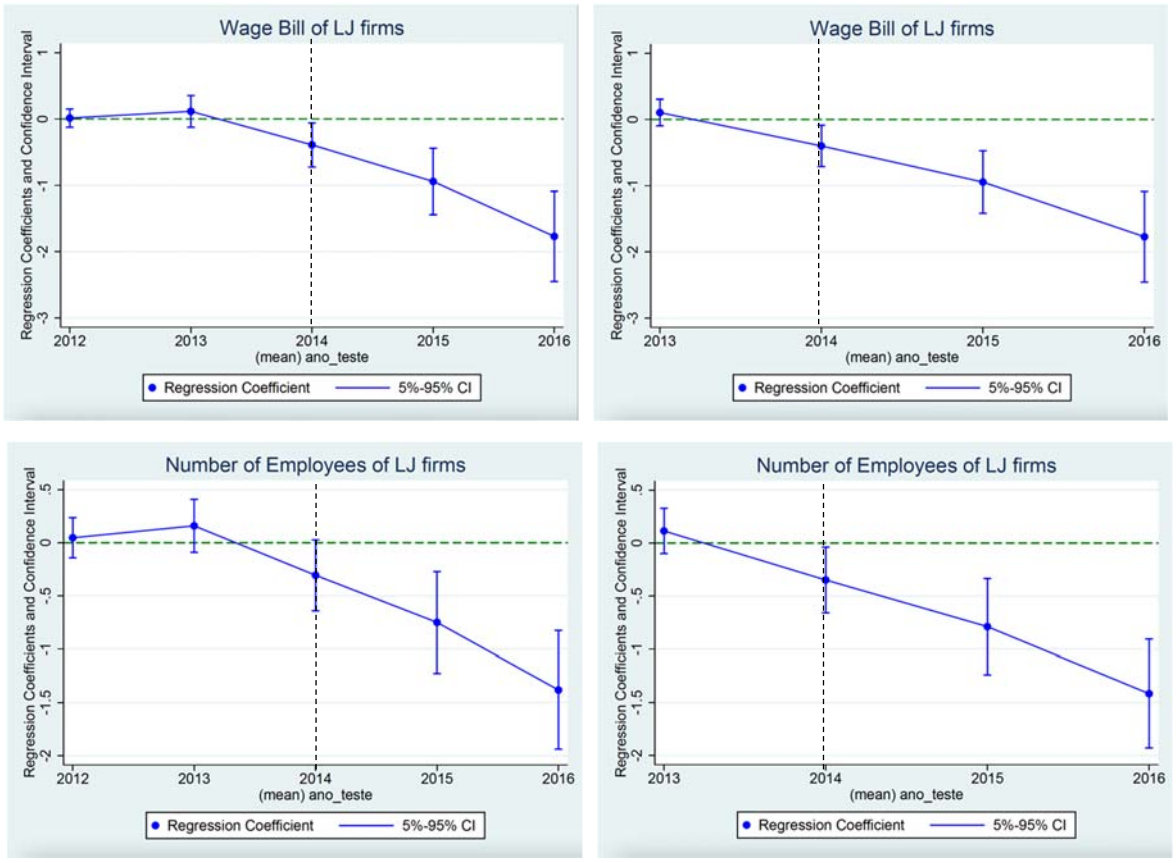
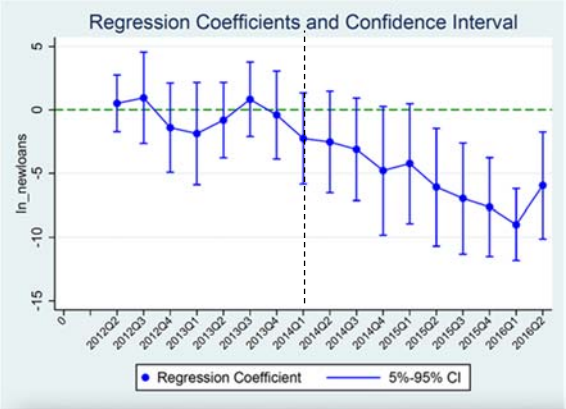


Figure 2: Effects on credit to Lava Jato firms

Panel A plots the corresponding DiD estimator with confidence intervals of the model $Ln(1+New\ Loans)_{i,t} = \alpha + \beta_1 Lava\ Jato_i + \beta_2 Post_t + \beta_3(Lava\ Jato_i \times Post_t) + \varepsilon_{i,t}$. This analysis considers all firms with more than 250 employees and includes the 21 Lava Jato firms. Post is decomposed in quarter dummies using 2012Q1 as reference category. Panel B displays new loans for Lava Jato firms and a synthetic control group.

Panel A: Volume of new loans (DiD estimator)



Panel B: Volume of new loans with synthetic control group

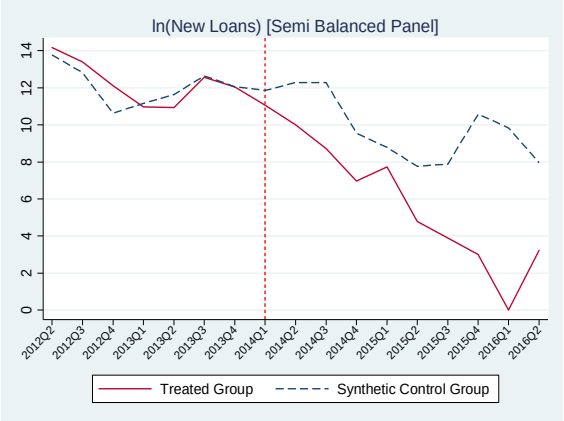


Figure 3: Bank exposure to Lava Jato firms

This figure shows the distribution of the variable *Lava Jato bank exposure*, which is the credit exposure of privately owned banks to Lava Jato firms in 2012 (bars), as well as the number of firm-quarter observations of each bank (lines). The vertical broken blue line indicates the tercile split of the sample (T3 vs. T2 and T1).

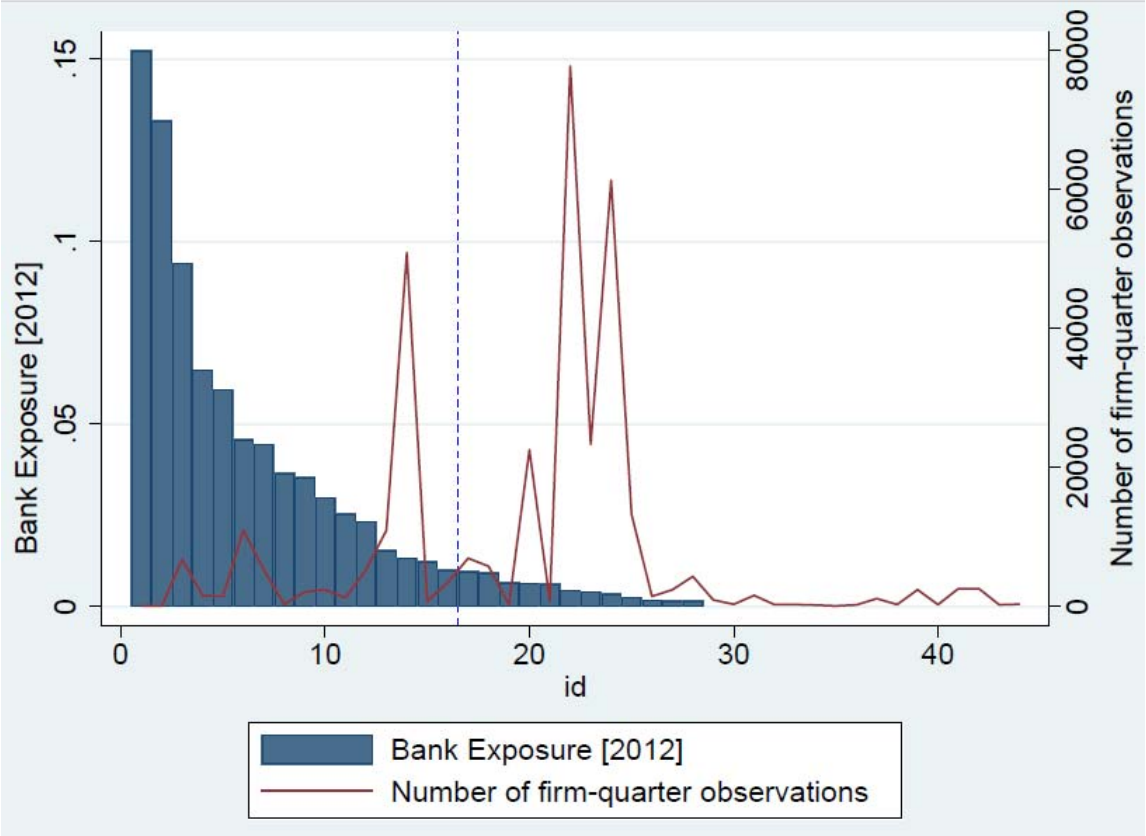


Table 1: Summary statistics

This table shows the summary statistics for the main variables used in the paper for the sample of 12,697 firms, 44 private banks and uses the period 2013Q1-2016Q2 (excluding information regarding the quarter 2014Q1). We excluded contracts with provisions above 5%, as well as negative interest rates from our sample. All the variables related to loan amount characteristics were winsorized at the 5% and 95% levels due to the presence of outliers in the original distribution.

Variable	Number of obs.	Mean	Median	Std. Dev.
Loan characteristics				
New Loans$_{it}$	17,809	20,541,316	756,420	88,038,727
New Loans$_{ibt}$	235,460	1,734,128	0	11,727,121
Maturity $_{ibt}$	80,339	11.1867	5	13.7284
Interest rate $_{ibt}$	60,332	47.0572	19.64239	82.2982
Regulatory Rating $_{ibt}$	80,339	2.2642	2	1.0272
Loan loss provision $_{ibt}$	80,339	.0089	.0049	.0107
Firm characteristics [2012]				
Number of employees$_i$	12,697	1,035.94	485	2,440.86
Wage bill$_i$	12,697	2174277	846464	6091897
Age$_i$	12,697	23.46	20.57	14.43
Election campaign donations$_i$	12,697	62.67	0	430.14
Number of bank relationships$_i$	12,697	3.52	3	3.10
Bank characteristics				
Liquidity$_{bt}$	510	.2212	.1980	.1391
Credit/Assets$_{bt}$	510	.5381	.5147	.2205
Log(Total Assets$_{bt}$*)*	510	23.37	23.10	1.56
Capital$_{bt}$	510	.18	.16	.08
NPL$_{bt}$	510	.0428	.0386	.0363
ROA$_{bt}$	510	.0077	.0101	.0199
Lava Jato exposures [2012]				
Lava Jato bank exposure$_b$	44	.0193	.0041	.0343
Lava Jato firm exposure$_i$	12,697	.0021	.0012	.0027

Table 2: Real effects on Lava Jato firms

Models (1) and (2) show the firm-level regression results of the model $C_{i,t} = \alpha + \beta_1 Lava\ Jato_i + \beta_2 Post_t + \beta_3 (Lava\ Jato_i \times Post_t) + \varepsilon_{i,t}$ where *C* indicates either *Wage bill or Number of employees*. *Lava Jato* is a dummy variable that equals one for the 21 construction firms that are subject to anti-corruption investigations and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Models (3) and (4) decompose the effect of the post period in 2014, 2015 and 2016. Standard errors (in parentheses) are clustered at the firm-level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

Dep. Var.	Log(wage bill) (1)	Log(employees) (2)	Log(wage bill) (3)	Log(employees) (4)
Lava Jato x Post	-1.123***	-0.947***		
	(0.188)	(0.172)		
Lava Jato x Post [2014]			-0.494***	-0.453***
			(0.127)	(0.124)
Lava Jato x Post [2015]			-1.029***	-0.884***
			(0.200)	(0.197)
Lava Jato x Post [2016]			-1.847***	-1.505***
			(0.328)	(0.247)
Firm-FE	Yes	Yes	Yes	Yes
Time-FE	Yes	Yes	Yes	Yes
Observations	48,037	48,455	48,037	48,455
Adj-R^2	0.024	0.104	0.026	0.105

Table 3: Effects on credit to Lava Jato firms

This table shows firm-level regression results of the model $C_{i,t} = \alpha + \beta_1 Lava\ Jato_i + \beta_2 Post_t + \beta_3 (Lava\ Jato_i \times Post_t) + \varepsilon_{i,t}$ where *C* indicates the volume of new loans *Ln(1+New loans)* or other loan characteristics (*Rating, Loan loss provision, Maturity* and *Interest rate*) from firm i at time t. This analysis considers all firms with more than 250 employees and includes the 21 construction firms cited by the Lava Jato investigations. The sample period of this analysis starts in the first quarter of 2012 and go until the second quarter of 2016. *Lava Jato* is a dummy variable that equals one for the 21 construction firms that are subject to anti-corruption investigations and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. The matching results reported in columns 2 and 3 consider the *nnmatch* algorithm proposed by Abadie et al. (2004) allowing the replacement of the selected units of comparison for either one corresponding match (1:1) (column 2) or four corresponding matches (1:4) (column 3). Robust standard errors (in parentheses) are clustered at the firm level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

Panel A: Volume of new loans

Dep. Var.:	*Log(1+New Loans)*		
	(1)	(2)	(3)
Lava Jato × Post	-5.164***	-3.760***	-3.827***
	(1.219)	(1.612)	(1.576)
Firm-FE	Yes	Yes	Yes
Time-FE	Yes	Yes	Yes
Matched sample (1:1)	No	Yes	No
Matched sample (1:4)	No	No	Yes
Number of observations	135,212	520	1,300
Adj-R^2	0.427	0.165	0.254

Panel B: Characteristics of new loans

Dep. Var.:	*Rating*	*Loan loss provision*	*Log(1+Maturity)*	*Interest rate*
	(1)	*(2)*	*(3)*	*(4)*
Lava Jato × Post	0.796***	0.006***	-0.097	4.917
	(0.197)	(0.002)	(0.273)	(8.506)
Firm-FE	Yes	Yes	Yes	Yes
Time-FE	Yes	Yes	Yes	Yes
Number of observations	65,353	65,353	65,353	54,512
Adj-R^2	0.447	0.415	0.445	0.418

Table 4: Effects on credit to non-investigated firms at the bank level

This table shows bank-level regression results of the model $Ln(1+New\ loans)_{j,t} = \alpha + \beta_1 Lava\ Jato\ bank\ exposure_j + \beta_2 Post_t + \beta_3 (Lava\ Jato\ bank\ exposure_j \times Post_t) + \varepsilon_{j,t}$ for bank j and time t. This analysis excludes Lava Jato firms. *Lava Jato bank exposure* is measured in different ways: in Column 1, it is a dummy variable that equals one if the bank is in the upper tercile of the distribution of the continuous exposure and zero otherwise, in Column 2 it is a continuous variable and in Column 3 shows the upper and mid-tercile vis-à-vis the bottom tercile as reference category. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

Dep. Var.:	*Log(New Loans)*		
	(1)	*(2)*	*(3)*
Lava Jato bank exposure (T3) × Post	-0.558***	-0.579***	
	(0.173)	(0.185)	
Lava Jato bank exposure (T2) × Post		-0.111	
		(0.137)	
Lava Jato bank exposure (cont.) × Post			-6.420**
			(3.184)
Bank controls$_{t-1}$	Yes	Yes	Yes
Bank-FE	Yes	Yes	Yes
Time-FE	Yes	Yes	Yes
Number of observations	511	511	511
Adj-R^2	0.140	0.141	0.129

Table 5: Effects on credit by borrower type at the bank level

This table shows the regression results of the model $C_{j,t} = \alpha + \beta_1 Lava\ Jato\ bank\ exposure_j + \beta_2 Post_t + \beta_3 (Lava\ Jato\ bank\ exposure_j \times Post_t) + \varepsilon_{j,t}$ where $C_{j,t}$ is the (log of) total new loans granted by bank j to either households (columns 1, 2 and 3) or to firms (columns 4, 5 and 6) . *Lava Jato bank exposure* is a dummy variable that equals one if the bank is in the upper tercile of the distribution of the continuous bank exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Households			Firms		
Dep. Var.:	$Log(New\ Loans)_{households}$			$Log(New\ Loans)_{firms}$		
Lava Jato bank exposure × Post	-0.022	0.096	-0.0129	-0.332*	-0.335*	-0.306*
	(0.169)	(0.284)	(0.167)	(0.171)	(0.186)	(0.165)
Lava Jato bank exposure		-0.642			0.243	
		(0.470)			(0.290)	
Post	0.071			0.126		
	(0.102)			(0.092)		
Bank controls $_{t-1}$	Yes	Yes	Yes	Yes	Yes	Yes
Bank-FE	Yes	-	-	Yes	-	-
Time-FE	-	Yes	-	-	Yes	-
Number of observations	427	428	427	536	536	536

Table 6: Effects on credit to non-investigated firms

This table shows firm-bank level regression results of the model $C_{i,j,t} = \alpha + \beta_1 Lava\ Jato\ Exposure_j + \beta_2 Post_t + \beta_3(Lava\ Jato\ bank\ exposure_j \times Post_t) + \varepsilon_{i,j,t}$ where $C_{i,j,t}$ denotes either New Loans [dummy] or Ln(1+New Loans). The analysis of the extensive margin considers only firms that did not borrow before 2014Q1, while the intensive margin considers only firms that did borrow before 2014Q1. Both analyses exclude Lava Jato firms. *Lava Jato bank exposure*, as defined in Table 1, is a dummy that equals one if the bank is in the upper tercile of the distribution of the continuous exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank-level. Standard errors (in parentheses) are clustered at the bank-time level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)
	Extensive margin		Intensive margin	
Dep. Var.:	*Dummy [New Loans>0]*	*Dummy [New Loans>0]*	*Log(1+New Loans)*	*Log(1+New Loans)*
Lava Jato bank exposure$_{T3}$ × Post	-0.010***	-0.012***	-0.105	-0.383*
	(0.002)	(0.003)	(0.175)	(0.211)
Bank controls $_{t-1}$	Yes	Yes	Yes	Yes
Firm-Time-FE, Bank-FE	Yes	No	Yes	No
Firm-Bank-FE, Time-FE	No	Yes	No	Yes
Number of observations	154,109	207,458	194,328	235,460
Adjusted-R^2	0.076	0.158	0.157	0.307

Table 7: Effects on credit to non-investigated firms moderated by firm election campaign donations

This table shows firm-bank level regression results of the model $C_{i,j,t} = \alpha + \ldots + \beta_4 Lava\ Jato\ bank\ exposure_j \times Post_t + \beta_5 Lava\ Jato\ bank\ exposure_j \times Campaign\ Donations_i + \beta_6 (Lava\ Jato\ bank\ exposure_j \times Campaign\ Donations_i \times Post_t) + \varepsilon_{i,j,t}$ where $C_{i,j,t}$ is *New Loans [dummy]*$_{i,j,t}$ for the extensive margin or *Ln(1+New Loans)*$_{i,j,t}$ for the intensive margin. Extensive margin considers only firms that did not borrow before 2014Q1, intensive margin considers only firms that did borrow before 2014Q1, both analyses exclude Lava Jato firms. *Lava Jato bank exposure* is a dummy that equals one if the bank is in the upper tercile of the distribution of the continuous measure of bank exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Electoral campaign donations is measured in 2011 and is represented by a dummy equal to one if the moderator is in the upper tercile of the distribution and zero otherwise.. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank-time level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Extensive margin			Intensive margin		
Dep. Var.:	*Dummy [New Loans >0]*			Log (1+New Loans)		
LJ exposure$_{T3}$ × *Post* × *Donations*	0.006*	0.007***	0.004	-0.428**	-0.428**	-0.353**
	(0.003)	(0.002)	(0.004)	(0.197)	(0.188)	(0.155)
LJ exposure$_{T3}$ × *Post*	-0.012***	-0.014***	-	0.009	-0.265	-
	(0.003)	(0.003)		(0.185)	(0.224)	
LJ exposure$_{T3}$ × *Donations*	-0.001**	-	0.000	0.546***	-	0.506***
	(0.000)		0.004	(0.172)		(0.118)
Post × *Donations*	-	-0.005***	-	-	-0.197**	-
		(0.001)			(0.095)	
Bank controls $_{t-1}$	Yes	Yes	No	Yes	Yes	No
Bank FE	Yes	-	-	Yes	-	-
Time FE	-	Yes	-	-	Yes	-
Firm-Time-FE,	Yes	-	Yes	Yes	-	Yes
Firm-Bank-FE	-	Yes	-	-	Yes	Yes
Bank-Time-FE	-	-	Yes	-	-	Yes
Number of observations	154,105	207,110	154,105	194,328	235,352	194,323
Adjusted-R^2	0.076	0.154	0.079	0.157	0.307	0.162

Table 8: Effects on credit to non-investigated firms moderated by number of bank relationships.

This table shows firm-bank level regression results of the model $C_{i,j,t} = \alpha + \ldots + \beta_4 Lava\ Jato\ bank\ exposure_j \times Post_t + \beta_5 Lava\ Jato\ bank\ exposure_j \times \#bank\ relationships_i + \beta_6 (Lava\ Jato\ bank\ exposure_j \times \#bank\ relationships_i \times Post_t) + \varepsilon_{i,j,t}$ where $C_{i,j,t}$ is *New Loans [dummy]*$_{i,j,t}$ for the extensive margin or *Ln(1+New Loans)*$_{i,j,t}$ for the intensive margin. Extensive margin considers only firms that did not borrow before 2014Q1, intensive margin considers only firms that did borrow before 2014Q1, both analyses exclude Lava Jato firms. *Lava Jato bank exposure* is a dummy that equals one if the bank is in the upper tercile of the distribution of the continuous measure of bank exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Number of bank relationships$_i$ is measured in 2012 and is represented by a dummy equal to one if the moderator is in the upper tercile of the distribution and zero otherwise. Bank controls are *Credit/Assets*, *Ln(Total Assets)*, *Capital*, *Non-performing loans*, *Liquidity* and *Return on Assets* lagged by one period. Standard errors (in parentheses) are clustered at the bank-time level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Extensive margin			Intensive margin		
Dep. Var.:	*Dummy[New Loans >0]*			Log(1+New Loans)		
LJ exposure$_{T3}$ × *Post* × *I* [*#bank relationships* >1]	-0.012***	-0.001	-0.018***	1.701***	1.479***	1.699**
	(0.004)	(0.003)	(0.006)	(0.652)	(0.487)	(0.732)
LJ exposure$_{T3}$ × *Post*	0.000	-0.013***	-	-1.790***	-1.816***	-
	(0.004)	(0.004)		(0.659)	(0.543)	
LJ exposure$_{T3}$ × *I* [*#bank relationships* >1]	-0.004***	-	0.000	-1.355**	-	-1.357**
	(0.001)		0.005	(0.571)		(0.596)
Post × *I* [*#bank relationships* >1]	-	0.007***	-	-	-0.011	-
		(0.002)			(0.165)	
Bank controls $_{t-1}$	Yes	Yes	No	Yes	Yes	No
Bank FE	Yes	-	-	Yes	-	-
Time FE	-	Yes	-	-	Yes	-
Firm-Time-FE,	Yes	-	Yes	Yes	-	Yes
Firm-Bank-FE	-	Yes	-	-	Yes	-
Bank-Time-FE	-	-	Yes	-	-	Yes
Number of observations	140,749	186,482	140,749	193,989	233,409	193,984
Adjusted-R^2	0.076	0.162	0.080	0.156	0.308	0.162

Table 9: Effects on credit to non-investigated firms

This table shows firm-level regression results of the model $C_{i,t} = \alpha + \beta_1 Lava\ Jato\ firm\ exposure_i + \beta_2 Post_t + \beta_3 (Lava\ Jato\ firm\ exposure_i \times Post_t) + \varepsilon_{i,t}$ where *C* indicates *Ln(1+New loans), log(New loans) (*where only positive values are considered), *Rating, Loan loss provision, log (1+Maturity)* and *Interest rate* from for i at time t. Apart from New loans which are the sum over all banks all other variables are weighted averages. This analysis considers all firms with more than 250 employees and includes the 21 construction firms cited by the Lava Jato investigations. The sample period of this analysis starts in the first quarter of 2012 and goes until the second quarter of 2016. *Lava Jato firm exposure* is a dummy that equals one if the firm is in the upper tercile of the distribution of the continuous indirect exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014Q1; the first quarter of 2014, in which the Operation Lava Jato started, is omitted. Robust standard errors (in parentheses) are clustered at the firm level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

Dep. Var.:	*Log(1+New Loans)*	*Log(New Loans)*	*Rating*	*Loan loss provision*	*Log(1+Maturity)*	*Interest rate*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
LJ firm exposure$_{T3}$						
$\times$ *Post*	-1.230***	-0.234***	0.118***	0.001***	0.016	0.882*
	(0.104)	(0.0343)	(0.019)	(0.000)	(0.018)	(0.520)
Firm-FE	Yes	Yes	Yes	Yes	Yes	Yes
Time-FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	150,608	73,270	73,270	73,270	73,270	61,625
Adj-R^2	0.050	0.008	0.020	0.034	0.011	0.061

Table 10: Real effects on non-investigated firms through firm exposure to Lava Jato

Models (1), (2), and (3) show firm-level regression results of the model $C_{i,t} = \alpha + \beta_1 Lava\ Jato\ firm\ exposure_i + \beta_2 Post_t + \beta_3 (Lava\ Jato\ exposure_i \times Post_t) + \varepsilon_{i,t}$ where *C* indicates either *Wage bill or Number of employees. Lava Jato firm exposure* is a dummy that equals one if the firm is in the upper tercile of the distribution of the continuous indirect exposure and zero otherwise. *Post* is a dummy variable that switches to one in the period after 2014. Models (3) and (4) decompose the effect of the post period in 2014, 2015 and 2016. Standard errors (in parentheses) are clustered at the firm-level. ***, **, * denote statistical significance at the 1%, 5% or 10% level.

Dep. Var.:	Log(Wage bill) (1)	Log(Employees) (2)	Log(Wage bill) (3)	Log(Employees) (4)
LJ firm exposure$_{T3}$ × Post	-0.109***	-0.0845***		
	(0.017)	(0.016)		
LJ firm exposure$_{T3}$ × Post [2014]			-0.071***	-0.054***
			(0.015)	(0.014)
LJ firm exposure$_{T3}$ × Post [2015]			-0.116***	-0.089***
			(0.020)	(0.018)
LJ firm exposure$_{T3}$ × Post [2016]			-0.142***	-0.112***
			(0.025)	(0.023)
Firm-FE	Yes	Yes	Yes	Yes
Time-FE	Yes	Yes	Yes	Yes
Number of observations	47,886	48,304	47,886	48,304
Adj-R^2	0.023	0.103	0.024	0.103