Motives and Consequences of Libor Strategic Reporting: How Much Can We Learn from Banks' Self-Reported Borrowing Rates?

Pietro Bonaldi*

July, 2023

Abstract

Libor is an estimate of interbank borrowing costs computed daily from rates reported by a fixed panel of banks. Evidence suggests that banks have manipulated Libor in recent years by misreporting their borrowing costs. I estimate a strategic reporting model that identifies banks' borrowing costs as well as their motives for misreporting. The estimation places a lower bound on the value that Libor would have had if banks had truthfully reported their borrowing costs. The model is identified even when unobserved heterogeneity exists in the form of a common cost component that is known by all banks but unobservable to the econometrician and is allowed to follow a nonstationary process. The only data used for identification are banks' Libor quotes. Overall, I find that the estimated lower bound for the unmanipulated Libor is always above the published Libor, with an average deviation of 25 basis points at the worst of the financial crisis of 2007 - 2008. The estimated bound displays a pattern similar to two other measures of interbank borrowing costs that have been used previously to assess the extent of manipulation. The model is also used to determine the extent to which misreporting was motivated by signaling or banks' net exposure to Libor. The estimation results indicate that sending creditworthiness signals was the main driver of systematic misreporting from 2007 to 2010.

*Carnegie Mellon Tepper School of Business (pbonaldi@andrew.cmu.edu). I would like to thank the editor, Áureo de Paula, and threee anonymous referees for their valuable feedback, which greatly improved this paper. Special thanks to my advisor, Ali Hortaçsu, the two other members of my thesis committee, Brent Hickman and Azeem Shaikh, and Jakub Kastl, for their advice, contributions and support. I am also grateful to Felipe García, for excellent research assistance, to Jorge Balat, Stephane Bonhomme, Juan Esteban Carranza, Mervyn King, Ignacia Mecadal, Juan Ospina, Harry J. Paarsch, Isabelle Perrigne, Philip Reny, Yuya Takahashi and Quang Vuong for very helpful comments, and to seminar and conference participants at Carnegie Mellon Tepper, Rice University, Queen's University, EIEF, Johns Hopkins Carey, University of Colorado Boulder Leeds, Duke Fuqua School of Business, Universidad del Rosario, European Central Bank, Bank of Canada, Banco de la República, Analysis Group, Cornerstone Research, Compass Lexicon, and the North American Summer Meeting of the Econometric Society 2016, for valuable feedback. A previous version of this paper was presented as my doctoral thesis at the University of Chicago, Department of Economics, under the title: "Libor Misreporting as a Bayesian Game with Unobserved Heterogeneity". All errors and omissions are my own.

1 Introduction

On April 16, 2008, the *Wall Street Journal* published an article suggesting that banks had been underreporting their borrowing costs by submitting quotes for the computation of Libor (London interbank offer rate) below estimates of such costs implied by credit default swaps (CDS) spreads (Mollenkamp and Whitehouse, 2008). Libor is a measure of the average cost of short-term, unsecured borrowing in the London interbank market. Before recent regulatory reforms, it was computed daily as a trimmed mean of interest rates self-reported by a fixed panel of large banks. In recent years, several banks have been investigated for their alleged attempts to manipulate Libor and other similar benchmarks. On May 21, 2015, the Council on Foreign Relations reported on its website that banks had paid more than $9 billion in fines and settlements with regulators for their involvement in the Libor scandal.[1] According to recent estimates reported by Schrimpf and Sushko (2019), approximately $400 trillion worth of financial contracts where indexed to Libor by mid-2018, although financial regulators worldwide had already started implementing strategies to transition from Libor to other benchmark rates.[2]

All banks in the Libor panel hold billions of dollars' worth of floating-rate assets and liabilities, and interest rate derivatives indexed to this benchmark. Consequently, they are exposed to variations in Libor. Libor is widely used as a measure of credit risk. Hence, a bank reporting above average Libor quotes could be seen as facing higher risks than others, which may further increase its funding costs or even trigger a run, as discussed by Duffie and Stein (2015).[3] Financial regulators, the academic literature and the press have all recognized that sending signals of creditworthiness to the market and benefiting from exposure to the benchmark were the two main drivers of Libor misreporting during the financial crisis (see Mollenkamp and Whitehouse, 2008; Duffie and Stein, 2015; Abrantes-Metz et al., 2012; and Gandhi et al., 2019).

In this study, I use structural econometric methods from the empirical auctions literature to estimate a model of strategic reporting that identifies a set of parameters determining banks' incentives for misreporting Libor, as well as the distributions of their true borrowing costs. Based on these estimates, I place a lower bound on the value Libor would have taken had banks truthfully reported their borrowing costs. The only data used in the main specification are the banks' reported quotes. The estimated bound is above the published Libor during the whole period considered, suggesting that Libor understated interbank borrowing costs, which is consistent with the results of previous studies. The model also identifies signaling and net exposures to Libor as distinct motives for mis-

[1] A more detailed account of the penalties that regulators imposed on banks can be found at: http://www.cfr.org/united-kingdom/understanding-libor-scandal/p28729

[2] Duffie and Stein (2015) provide a detailed account of the crucial role that Libor played for many years in financial markets globally.

[3] On June 7, 2018 the Financial Times published a report on previously undisclosed court documents showing that the Department of Justice had evidence in 2015 suggesting that high-level executives at Société Générale ordered the bank's attempt to manipulate the USD Libor. As described in the report, "in a meeting on May 21 2010, as European banks struggled with the eurozone crisis, the first executive told a manager "to lower Société Générale's Libor submissions" after complaining that they were higher than other banks, according to the DoJ." (Shubber and Keohane, 2018)

reporting. The results indicate that the main driver of misreporting during the financial crisis was an attempt to send creditworthiness signals to the market.

In the model, banks play a Bayesian game. Each day, each bank receives a private shock to its borrowing cost in the interbank market and reports a quote to the regulator. Banks interact strategically, they all potentially benefit from manipulating the benchmark rate, given their individual net exposure to it, and from underreporting their private borrowing costs to signal themselves as more creditworthy than their peers. They also face misreporting costs due to potential loss of reputation or regulatory penalties. Banks make their strategic decisions both simultaneously and independently.

Snider and Youle (2012) propose a Libor reporting game with complete information and find evidence of manipulation driven by exposure to Libor; their model does not include signaling. Chen (2013) assumes symmetric independent private costs and adds signaling to banks' objective functions. He proves the existence of a Bayesian Nash equilibrium in pure non-decreasing strategies. I extend his results and show that equilibrium strategies are strictly increasing. This result is necessary for my identification strategy, which is based on inverting the equilibrium strategies, following an idea from the auctions literature originally introduced by Guerre et al. (2000). Youle (2014) estimates the parameters of the model capturing exposure-based incentives to manipulate Libor. His identification strategy relies mainly on additional data sources, such as CDS spreads, to approximate banks' borrowing costs and does not address unobserved heterogeneity. It does not identify the signaling motives for misreporting and, thus, cannot recover the unmanipulated Libor.

My model coincides with Youle (2014) in the specification of banks' payoff function. Nonetheless, crucial differences exist in the way I model the process followed by borrowing costs, which are necessary for a valid identification and estimation of the model primitives. I assume that daily borrowing costs include a time-varying common component that is known by all banks, but not observable to the econometrician, which allows for unobserved heterogeneity. I contribute to the structural estimation literature by allowing the common cost component to follow a nonstationary process, relaxing an assumption in Li et al. (2000) and Krasnokutskaya (2011). In the present context, assuming either time-independence or stationarity likely results in largely inaccurate estimates, given the high persistence of interest rates (Rose, 1988). Instead, I propose an extension of the deconvolution method of Li and Vuong (1998), Li et al. (2000) and Krasnokutskaya (2011) that removes the common component.[4] This allows me to identify the model parameters and distributions from data on Libor quotes alone, without relying on any other information on banks' borrowing costs.

Removing the common cost component identifies the distributions of normalized quotes, defined as the quotes that banks would have reported if there were no variation in common costs, but only up to an unknown additive constant. To identify a lower bound on the unmanipulated Libor, I use the lowest signaling motives that are consistent with the model estimates. Higher values for the

[4] Li et al. (2000) and Krasnokutskaya (2011), among others, have studied unobserved heterogeneity that arises in the estimation of procurement auctions when bidders share common costs that are unobservable to the econometrician.

signaling parameters would imply a higher unmanipulated Libor because sending creditworthiness signals gives banks incentives to understate their borrowing costs.

The identification of signaling and exposure to Libor as distinct drivers of misreporting relies on two assumptions. First, I assume that each bank's idiosyncratic borrowing costs are drawn daily from a distribution with the same mean and median, which is weaker than assuming that these distributions are symmetric. Roughly speaking, symmetry would imply that on any given day banks are equally likely to receive a positive or a negative idiosyncratic shock to their mean borrowing cost. Second, I assume that the means are the same for all banks in the US Dollar Libor panel, but I allow their variances and all other higher moments to differ across banks. For each bank, the idiosyncratic private cost is the daily difference between the interest rate it pays for an average-size loan from other bank and the common cost. In turn, the latter is the mean interbank rate. Hence, the second assumption requires that the borrowing costs of no bank in the panel deviate substantially, on average across several days, from the overall interbank market mean. With truthful reporting, the distributions of normalized quotes would coincide with the distributions of idiosyncratic costs. However, the estimated normalized quotes do not have the same means across banks and, for a given bank, do not generally have the same mean and median. In the model, these two differences are explained by two parameters: signaling and net exposure to Libor. As a robustness check, I estimate an alternative model that drops the assumption that the mean idiosyncratic costs are the same for all banks. The results regarding the relative importance of signaling and exposure to Libor do not seem to depend on these assumption. However, this version of the model does not provide an estimate of the lower bound of the unmanipulated Libor. Instead, it requires an observable measure of the common cost to identify signaling and exposure separately. Therefore, I maintain this assumption in the baseline model.

I estimate the model using data on the USD three-month Libor from September 9, 2007 to December 31, 2010. The estimated lower bound of the counterfactual unmanipulated Libor is always above the published Libor. In the aftermath of Lehman Brother's bankruptcy, the lower bound is, on average, 25 basis points (0.25 percentage points) higher than the published Libor. The lower bound on the unmanipulated Libor follows similar patterns than two other measures of average costs for unsecured funding described by Kuo, Skeie, and Vickery (2012). All three metrics indicate that Libor understated borrowing costs during this period. Regarding banks' incentives to misreport, the estimates strongly suggest that the deviation between Libor submissions and borrowing costs is mostly explained by banks' attempts to send signals of creditworthiness. According to these estimates, previous studies focused exclusively on exposure may have overlooked the main driver of misreporting. The results also show large changes over time in the signaling incentives to misreport, and suggest that these increased dramatically immediately after Lehman Brother's bankruptcy. A likely explanation is that banks' signaling incentives responded to a sharp rise in the market's perception of default risk, as measured by CDS spreads. Indeed, a simple regression of the estimated signaling parameters on CDS spreads confirms a positive association.

Previous empirical work by Abrantes-Metz, Villas-Boas, and Judge (2011) and Abrantes-Metz, Kraten, Metz, and Seow (2012) suggests that Libor misreporting was a generalized practice, not restricted to a few banks or dates. Cassola, Hortaçsu, and Kastl (2013) estimate the maximum rates that banks are willing to pay for short-term collateralized loans at liquidity auctions conducted by the European Central Bank. They find that Euribor[5] quotes are usually below the estimated willingness to pay for secured loans during 2007 and 2008, even though these are quotes for unsecured borrowing, and thus should include a risk premium. Similarly, Kuo, Skeie, and Vickery (2012) compare Libor quotes to other measures of individual borrowing costs from 2007 to 2009 and find evidence of underreporting. They also analyze changes in the spreads between Libor and other estimates of average costs for unsecured funding, specifically, the New York Funding Rate (NYFR) and the Eurodollar deposits rate (EDR) published by the Federal Reserve, showing that during the period considered the spread is negative and wider than in previous years, which suggests that Libor understated borrowing costs during the financial crisis. Gandhi et al. (2019) find that banks' excess equity returns correlate with changes in Libor and with their individual reports. They interpret this as evidence of exposure and signaling incentives to manipulate Libor. By contrast, I find no evidence of manipulation before September 2007, when the signaling incentive was presumably negligible. This finding is broadly consistent with the behavior of other funding costs measures. During a period of two years prior to September 2007, the spread between the EDR and Libor is mostly positive and much smaller in magnitude than in the following three years.

The remainder of this paper is organized as follows. Section 2 describes some measures of banks' borrowing costs and credit risk that are indicative of Libor underreporting. Section 3 presents a model of strategic Libor reporting. Section 4 discusses the identification strategy. Section 5 describes the estimation procedure. Section 6 presents estimates of the extent of misreporting separately attributable to signaling and exposure to Libor. It also compares Libor to an estimated lower bound on the unmanipulated Libor. In Section 7 I discuss a counterfactual analysis that assesses the effectiveness of current proposals to improve interest rate benchmarks.

2 Background

Libor is an average of interbank offer rates, computed daily for different currencies and maturities, based entirely on quotes reported by a panel of banks. Each day, each bank in the panel reports a rate answering the question "At what rate could you borrow funds, were you to do so by asking for and then accepting interbank offers in a reasonable market size just prior to 11 am?" After receiving all quotes, the regulator publishes the benchmark rate and also the set of all quotes submitted by the banks with their respective identities.[6] The benchmark is a trimmed mean of these quotes after dropping roughly the 25% highest and 25% lowest quotes. For each currency, the panel consists of

[5]Euribor stands for Euro Interbank Offered Rate. It is similar to Libor but for the euro interbank market.

[6]From 1985 to 2013, the British Banking Authority (BBA) was responsible for the administration of Libor. As a consequence of the manipulation scandal, on February 1 2014, administration was handed over to the Intercontinental Exchange Benchmark Administration (IBA). Since the scandal broke, several changes have been implemented following recommendations by Wheatley (2012). Starting in April 2013, individual quotes are now published with a lag of three months after the submission date.

approximately 16 leading banks trading in London that borrow and lend in this currency.[7]

For at least two years in the midst of the financial crisis, Libor was consistently below other measures of borrowing costs for unsecured funds in the interbank market. Two such measures are the NYFR and the EDR. The NYFR was based on a survey of contributors that trade in interbank markets. In contrast to Libor, individual quotes were not published and the identity of the contributors was unknown. The EDR is computed from quotes provided by brokers who act as intermediaries in the interbank market (see Kuo et al. (2012) for a more detailed description of these rates and how they compare to Libor). Figure (1) shows the spreads between Libor and each of these rates. Libor followed the EDR closely since its inception until late 2007, mostly lying a few basis points above it (1 basis point is equal to 0.01 percentage points; henceforth I use the abbreviation "bp"). The Libor-EDR spread becomes negative and relatively large in magnitude around August 2007, when BNP Paribas announced that it had suspended redemption in two subprime mortgage funds. It then drops drastically with Lehman's collapse reaching its lowest historical value (approx. -200 bp). As the crisis evolves, this spread narrows but remains negative and large in magnitude compared to its pre-crisis levels. The NYFR was introduced on June 2, 2008, and was discontinued in 2012. For a few months before Lehman Brothers' bankruptcy, the Libor-NYFR spread was, on average, no lower than -1.5 bp. However, this difference increased substantially in magnitude after Lehman's collapse and remained large and negative for at least one year.

CDS spreads on bank bonds increased dramatically during the crisis raising concerns that Libor quotes were not accurately capturing the underlying credit risks (Mollenkamp and Whitehouse, 2008). CDS spreads are a form of fee protection buyers pay to insure against default or other credit events. Hence, they are measures of the perceived likelihood of a credit event, and ultimately of perceived credit risk.[8] Figure 1 shows the mean spread for five-year CDS on bonds issued by banks in the USD Libor panel. Clearly, the financial crisis and subsequent sovereign debt crisis in Europe were accompanied by a sharp and persistent increase in banks' CDS spreads. Bank executives may have felt pressured to use Libor quotes as signals of creditworthiness. The results in Section 6.3 indicate that higher CDS spreads are associated with higher signaling motives.

Finally, as the crisis unfolded, market participants and regulators became increasingly concerned that Libor reports were not accurately reflecting borrowing costs. This may have raised the expected costs of misreporting Libor due to potential loss of reputation or regulatory penalties, and at least partially counteracted the incentives to misreport.

3 The Model

Banks choose Libor quotes in a complex dynamic environment where they borrow from each other, take short and long positions indexed to Libor, and report these quotes knowing that they will

[7]There is some variation in the composition of Libor panels across currencies and over time. In particular, there were sixteen banks in the USD Libor panel from September 2007 to December 2010, the period I use for the estimation of the model, although there was a change in its composition in January 2009.

[8]Blanco et al. (2005) provide empirical evidence that CDS spreads are useful indicators of credit risk. Cassola et al. (2013) find a positive correlation between five-year CDS spreads and banks' short term funding costs.

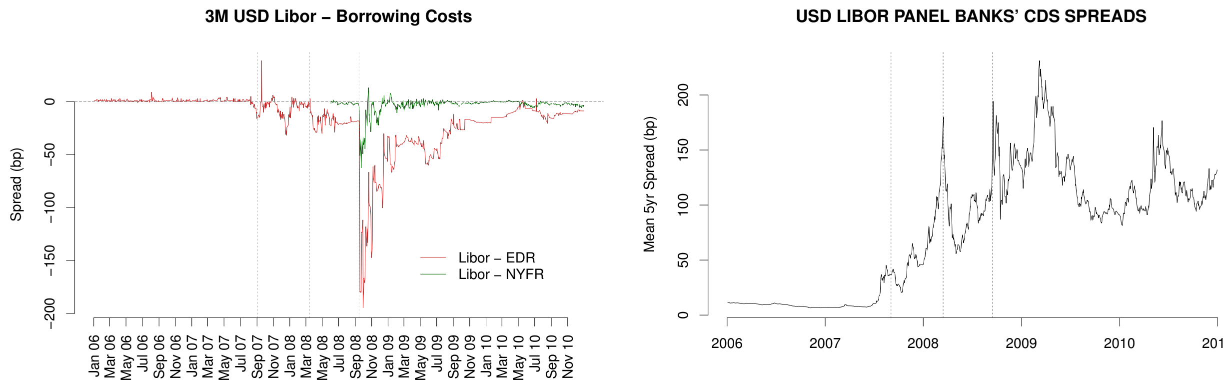

Figure 1: Borrowing costs and CDS spreads

The left panel shows the spread between the three-month USD Libor and the Eurodollar deposit rate (EDR) or the New York Funding Rate (NYFR). The right panel shows the average spread of five-year credit default swap (CDS) on bonds issued by banks in the Libor panel.The EDR is a measure of Eurodollar deposits based on quotes from brokers with access to data on interbank transactions. The NYFR is based on a survey of contributors that participate in the market for unsecured funds. CDS are measures of perceived credit risk in bond markets. Three dates are marked with vertical lines. (a) 09/03/2007: First trading day after BNP Paribas suspended redemption in two subprime mortgages funds. (b) 03/16/2008: JPMorgan Chase acquired Bear Stearns for less than 7% of its market value two days before (c) 09/15/2008: Lehman Brothers filed for bankruptcy protection. All spreads are measured in basis points (bp).

be publicly available soon after submission. The model below focuses exclusively on the reporting game, taking as given the equilibria in the interbank market and the markets for Libor-indexed loans and derivatives. Conditional on their behavior in these markets, banks realize there are potential gains and losses from reporting rates different than their borrowing costs. Since these gains depend at least partially on the benchmark rate, banks interact strategically when reporting their rates.

There is a fixed set of banks, $\mathcal{N} = \{1, ..., N\}$, that participate in the game. Each day t, within period p, bank $i \in \mathcal{N}$ observes its own borrowing costs in the interbank market for unsecured funds, $s_{i,t,p}$. The regulator asks all banks to report their costs, and in response, each submits a quote $r_{i,t,p}$ chosen from a set of admissible reports $\left[\underline{r}_p, \bar{r}_p\right]$. For all $i \neq j$, $s_{i,t,p}$ and $s_{j,t,p}$ are assumed to be independent, conditional on a common cost component $c_{t,p}$ that is common knowledge. The distributions of costs might differ for different banks, and are also common knowledge. A period corresponds to a time window of several days, within which some parameters of the model are assumed to remain constant.

3.1 Banks' Incentives

Banks submit their quotes simultaneously to maximize their expected payoff from participating in the static game, conditional on $c_{t,p}$. After observing all the quotes, the regulator computes the reference rate $\tilde{r}_{t,p}\left(r_{1,t,p}, ..., r_{N,t,p}\right)$, and reveals $\tilde{r}_{t,p}$ and the whole vector of individual quotes to the public. Banks could benefit from the game in two different ways: by manipulating the reference rate (in a direction that is consistent with their exposure) and by sending signals to the market of their creditworthiness. Correspondingly, their payoff functions include a linear gain on the level of the reference rate $\tilde{r}_{t,p}$ and another term that is linear on the difference between the reference rate and their quote. Additionally, a cost component that is quadratic in the difference $s_{i,t,p} - r_{i,t,p}$ reflects expected costs from reporting rates different from the underlying costs owing to potential credibility loss or regulatory penalties. Formally, bank i's expected payoff from reporting $r_{i,t,p}$ after observing both $s_{i,t,p}$ and $c_{t,p}$ is

$$u_i\left(r_{i,t,p}; s_{i,t,p}, c_{t,p}\right) = E_{R_{-i,t,p}}\left[\underbrace{\alpha_{i,p}\tilde{r}_{t,p}\left(r_{i,t,p}, R_{-i,t,p}\right)}_{\text{exposure}} + \underbrace{v_p\left(\tilde{r}_t\left(r_{i,t,p}, R_{-i,t,p}\right) - r_{i,t,p}\right)}_{\text{signaling}} - \underbrace{\gamma_p\left(s_{i,t,p} - r_{i,t,p}\right)^2}_{\text{cost}} | c_{t,p}, s_{i,t,p}\right] \quad (1)$$

where $R_{-i,t,p}$ is the vector of all other banks' reported rates, which is a random variable for bank i.[9] By assumption, $v_p > 0$ and $\gamma_p > 0$, but $\alpha_{i,p}$ is allowed to have an arbitrary sign reflecting the possibility that banks' exposure to the benchmark might be long or short. All these constants are common knowledge thus the only source of uncertainty for any bank is other banks' reports. The reference rate $\tilde{r}_{t,p}$ is a function of the vector of quotes submitted by all banks and is defined by:

$$\tilde{r}_{t,p} = \frac{1}{\tilde{N}} \sum_{k=\underline{n}+1}^{\bar{n}-1} r_{t,p}^{(k)} \quad (2)$$

where $r_{t,p}^{(k)}$ is the k-th smallest element of the vector of quotes $(r_{1,t,p}, ..., r_{N,t,p})$. $\underline{n}$ and $\bar{n}$ are cut-offs set by the regulator, and known by the banks, determining which quotes are trimmed; and $\tilde{N}$ is the effective number of quotes used to compute the trimmed mean.

Assuming a specific functional form for banks' payoff might raise some concerns on how accurately the model captures their incentives. Although there is a trade-off between generality and tractability, Equation (1) may still approximate banks' incentives reasonably well. The linear term $\alpha_{i,p}\tilde{r}_{t,p}$ can be interpreted as the periodic cash flow of bank i from its net exposure to Libor. As long as the exposure remains constant, the change in this cash flow due to an increase in $\tilde{r}_{t,p}$ of 1 bp should also be constant, regardless of the level of the reference rate. Moreover, a single bank cannot cause the reference rate to differ by a large margin from the value it would have under truthful reporting, since extreme quotes are not included in the computation of the trimmed mean. Hence, we should not expect a bank to deviate significantly from truthful reporting as long as there is an expected reputational or regulatory cost from doing so. The signaling and cost components of the payoff function can be interpreted as first and second order approximations to a more general specification of the respective gains and costs. Higher order effects may be present, but they should not be the main drivers of banks' quotes.

The signaling component captures the reputational gain that bank i derives from reporting a low borrowing rate $r_{i,t}$, relative to other banks, as described in Section 2. It is broadly consistent with descriptions of signaling incentives in previous academic literature on Libor manipulation and with anecdotal evidence reported in the financial press.

Regarding costs, the quadratic form $\gamma_p\left(s_{i,t,p} - r_{i,t,p}\right)^2$ is meant to bind the extent of misreporting

[9] I follow the convention of using uppercase letters for random variables, and lowercase letters for their realizations.

by, presumably, capturing expected costs due to loss of credibility and likely regulatory penalties.[10] The functional form chosen is the lowest-order polynomial approximation that is symmetric around zero ($s_{i,t,p} - r_{i,t,p} = 0$) and also attains its minimum at zero (hence the linear term is not included). It is symmetric because, ex-ante, banks are similarly likely to have incentives to over or understate their borrowing costs, and deviations in both directions are potentially damaging for other market participants who borrow and lend at rates indexed to Libor.

Banks in the model behave competitively and choose their reports privately to maximize their individual payoffs. They rely on information about each other incentives to devise their optimal strategies. The model does not allow for explicit collusion as a source of systematic manipulation for the following reasons. For each bank, its net exposure to Libor is the result of thousands of loans and derivatives. Given that the amounts at stake are substantial, it is reasonable to expect profit maximizing banks to choose their Libor quotes based on their net exposure, rather than on the specific positions of individual traders. There is indeed evidence that top bank executives were involved in manipulation attempts (see Section 2). Moreover, banks with opposing net exposures have strong incentives to compete rather than cooperate. However, it is presumably much harder for managers to collude than for individual traders due to much stricter regulatory scrutiny, even more so if it requires them to coordinate on both their Libor submissions and their net exposure to Libor.

The vast majority of fines and settlements that banks have paid due to their involvement in the manipulation of the USD Libor are related to individual rather than concerted wrongdoings. However, regulators in Europe have imposed fines on several banks after presenting evidence that some traders participated in cartels to try to manipulate Euribor and the Japanese Yen Libor. In consequence, in spite of the previous arguments, I cannot fully rule out the existence of similar cartels in the USD Libor panel. It is hard to anticipate, though, how ignoring such coalitions, if they are present, affects my estimates as any potential bias would depend on the number of coalitions and their respective incentives. For instance, two coalitions with opposing net exposures to Libor would likely compete in equilibrium to push the benchmark rate in different directions. The identification strategy in Section 4 should still extract information on the sign of their exposures from the distributions of their quotes, but the estimated magnitudes may be biased if the model ignores that banks internalize the positive externality on their coalition.

I will now introduce a set of assumptions that I use to characterize the Bayesian Nash Equilibrium of the non-cooperative game.

Assumption 1. *For all $i \in \mathcal{N}$, at each period p, the parameters $\alpha_{i,p}$, $v_{i,p}$ and $\gamma_{i,p}$ are common knowledge.*

A potential concern with Assumption 1 is that banks may not know each other's exposure to Libor.[11] However, such detailed knowledge of other banks' balance sheets is not necessary for banks

[10] As a matter of fact, almost all banks in the USD Libor panel have paid fines and settlements due to their involvement in the manipulation scandal.

[11] A weaker version of Assumption 1, where $(\alpha_{i,p}, v_{i,p}, \gamma_{i,p})$ is i's private information, but it is drawn from a

to choose their optimal strategies. As shown below, the first order condition of (1) implies that a weaker sufficient condition is that in equilibrium banks form correct beliefs about the distribution of reported rates. In a Bayesian fashion, by playing the game repeatedly and observing the actions of others after each instance of the game, together with other information they gather from trading with each other in the interbank market, banks would likely be able to correctly infer this distribution and converge to the equilibrium implied by Assumption 2.

3.2 Borrowing Costs

The cost of each bank ($s_{i,t}$) is determined by a common component c_t, known to all banks, and a private shock $\pi_{i,t,p}$. Each day t, $\Pi_{i,t,p}$ is drawn independently from the same distribution $F_{i,p}$.

Assumption 2. *For all* $i \in \mathcal{N}$ *and all* (t, p), $s_{i,t,p} = c_{t,p} + \pi_{i,t,p}$. $c_{t,p}$ *is common knowledge but* $\pi_{i,t,p}$ *is only privately known.*

In their work on the design of robust interest rates benchmarks, Duffie and Dworczak (2021) also assume that private borrowing costs are the sum of a common component and a zero mean idiosyncratic shock. Unlike them, I also assume that $c_{t,p}$ is common knowledge, but not observed by the econometrician. Cassola et al. (2013) study liquidity auctions in which banks periodically borrowed short-term funds from the European Central Bank in 2007 and 2008. Such funds were presumably close substitutes for interbank loans, and the set of bidders contains all banks in the USD Libor panel. They assume independent private values, conditional on commonly observed information. Assumption 2 is a particular case, with additively separable private and common components. I interpret the empirical results of Cassola et al. (2013) as a validation of their assumption.

Assumption 3. *For all* i, $\Pi_{i,t,p} \overset{iid}{\sim} F_{i,p}$, *and for any* (t, p) *and* $j \neq i$, $\Pi_{i,t,p}$ *and* $\Pi_{j,t,p}$ *are independent. The support of* $F_{i,p}$ *is a closed interval* $\left[\underline{\pi}_{i,p}, \bar{\pi}_{i,p}\right]$ *on the real line. The distributions* $F_{i,p}$, *for all* i, *are also common knowledge.*

Note that the independence of $S_{i,t,p}$ and $S_{j,t,p}$, conditional on $c_{t,p}$, follows directly from Assumptions 2 and 3. In addition, from Assumption 3, let $\mathcal{S}_{i,p}(c) = \left[c + \underline{\pi}_{i,p}, c + \bar{\pi}_{i,p}\right]$ be the support of $s_{i,t,p}$, conditional on c. $\mathcal{S}_{i,p}(c)$ is the set of all possible types for each player i, given c, at period p.

Assumption 4. *For all* i, *the random variable* $\Pi_{i,t,p}$ *is absolutely continuous with respect to the Lebesgue measure.*

The last assumption guarantees that both $\Pi_{i,t,p}$ and $\Pi_{i,t,p}$ have Lebesgue integrable densities on their support.

commonly known distribution for all $i \in \mathcal{N}$, still implies a model that has a Nash equilibrium in pure monotone strategies. McAdams (2003) proves existence in pure monotone strategies for a broad class of games with multi-dimensional types that includes this one as a particular case.

3.3 Bayesian Nash Equilibrium

The focus of this section is to describe the Bayesian Nash equilibrium (BNE) of the static game that banks play every day (t, p), hence I drop indices t and p for convenience. I first use results from previous literature to show that there is a BNE in pure, non-decreasing strategies. Then, I show that these strategies are strictly increasing in a subset of their domain. These results are crucial since the identification strategy in Section 4 is based on the existence of strictly increasing inverse equilibrium strategies.

From Equation (3), bank i's best response is given by

$$\varrho_i(s_i) = \underset{r_i \in [\underline{r}, \bar{r}]}{\operatorname{argmax}} E_{R_{-i}} \left[\alpha_i \tilde{r}(r_i, R_{-i}) + v_i (\tilde{r}(r_i, R_{-i}) - r_i) - \gamma_i (s_i - r_i)^2 | c, s_i \right] \tag{3}$$

for all $s_i \in \mathcal{S}_i(c)$. Note that banks are restricted to choose their quotes from a closed interval $[\underline{r}, \bar{r}]$. For instance, they may not be allowed to report borrowing rates below zero.

The maximization problem in (3) could have more than one solution for some s_i, hence ϱ_i may be set-valued or, equivalently, a correspondence. Since I am interested in single-valued equilibrium strategies, I use ρ_i to denote a single-valued function such that $\rho_i(s_i) \in \varrho_i(s_i)$, for all $s_i \in \mathcal{S}_i(c)$.

The payoff function u_i satisfies the single crossing property (SCP) of Milgrom and Shannon (1994), for any set of strategies ρ_{-i}, as can be verified by direct differentiation (in fact, $\frac{\partial^2 u_i}{\partial s_i \partial r_i} = 2\gamma_i > 0$). Therefore, the best response correspondence ϱ_i is non-decreasing in the strong set order (see Lemma 1 in Athey, 2001),[12] and it satisfies the *single crossing condition for games of incomplete information* of Athey (2001). It follows from her Corollary 2.1 that the game has a Bayesian Nash equilibrium in pure non-decreasing single-valued strategies. Moreover, since the best response ρ_i is non-decreasing, regardless of the strategies used by all other players, it follows that in any pure strategy Bayesian Nash equilibrium of this game all strategies are non-decreasing.

A set of necessary conditions for $(\rho_1, ..., \rho_N)$ to be a vector of equilibrium strategies can be obtained from the first order conditions of the payoff maximization problem (3). For any $i \in \mathcal{N}$ and $s_i \in \mathcal{S}_i(c)$, if $\rho_i(s_i)$ is an interior solution to problem (3), when all other banks follow strategies ρ_{-i}, then

$$(\alpha_i + v_i) \frac{\partial E_{S_{-i}} [\tilde{r}(\rho_i(s_i), \rho_{-i}(S_{-i})) | c, s_i]}{\partial r_i} - v_i + 2\gamma_i (s_i - \rho_i(s_i)) = 0 \tag{4}$$

where $\rho_{-i}(S_{-i}) = \{\rho_j(S_j) : j \neq i\}$ and $E_{S_{-i}}[\cdot | c, s_i]$ denotes the conditional expectation operator *from the perspective of bank* i, that is, when s_i is known, but for all $j \neq i$, S_j is a random variable with known conditional distribution.

Since ρ_j is strictly monotone, $E_{S_{-i}}[\tilde{r}(\rho_i(s_i), \rho_{-i}(S_{-i})) | c, s_i] = E_{R_{-i}}[\tilde{r}(r_i, R_{-i}) | c, s_i]$. Notice that Assumptions 2 and 3 imply that $E_{R_{-i}}[\tilde{r}(r_i, R_{-i}) | c, s_i] = E_{R_{-i}}[\tilde{r}(r_i, R_{-i}) | c]$. That is, given knowledge of c, its own cost s_i does not provide any further information to bank i regarding the distribution of other banks' quotes (costs).

[12] In Appendix A, I show that, as a correspondence, ϱ_i is non-decreasing in a very strong sense, i.e, for all $s, s' \in \mathcal{S}_i(c)$, with $s < s'$, $\max \varrho_i(s_i) \leq \min \varrho_i(s_i')$. By Berge's Theorem of the Maximum, we know ϱ_i is compact valued, so each of these sets has a minimum and a maximum.

As noted by Chen (2013), the derivative $\frac{\partial E_{R_{-i}}[\tilde{r}(r_i,R_{-i})|c]}{\partial r_i}$ can be expressed in terms of the probability that bank i assigns to the event that its quote would be used in the calculation of the benchmark rate, that is, the probability that $R^{(\underline{n})} \leq r_i \leq R^{(\bar{n})}$ (see Appendix A for a proof of this statement). More precisely,assumption that the mean idiosyncratic costs are the same for all banks

$$\frac{\partial E_{R_{-i}}\left[\tilde{r}\left(r_i, R_{-i}\right)|c\right]}{\partial r_i} = \frac{1}{\tilde{N}} P_{R_{-i}} \left\{ R^{(\underline{n})} \leq r_i \leq R^{(\bar{n})} | c \right\} \tag{5}$$

where the probability is consistent with the equilibrium distribution of R_{-i}. Let $\phi_i\left(r_i|c\right)$ denote such probability. Notice that it depends on the vector of equilibrium strategies of all other banks and the distributions of their costs. In Appendix A, I derive an explicit expression for $\phi_i\left(r_i|c\right)$ in terms of the distributions of borrowing costs.

Since α_i, v_i and γ_i cannot be separately identified, without loss of generality, I normalize the payoff function by dividing it by $2\gamma_i$. Henceforth, α_i and v_i denote the potential gains from exposure and signaling expressed as a fraction of the misreporting costs. Letting $\beta_i = \frac{\alpha_i + v_i}{\tilde{N}}$, the expression for the necessary condition in (4) simplifies to:

$$\beta_i \phi_i\left(\rho_i\left(s_i\right)|c\right) - v_i + \left(s_i - \rho_i\left(s_i\right)\right) = 0 \tag{6}$$

For each bank i, its equilibrium strategy ρ_i satisfies Equation (6), for all $s_i \in \mathcal{S}_i\left(c\right)$ such that $\rho_i\left(s_i\right)$ is an interior solution to problem (3). A natural lower bound for the set of admissible quotes is $\underline{r} = 0$. However, there is no obvious restriction on how large the reported rates should be allowed to be. I show in Appendix A that the upper bound $\bar{r}$ can be chosen in such a way that it is never binding. It follows that any positive optimal quote r_i^* is an interior solution, therefore, for all such $r_i^* > 0$ there is a s_i such that:

$$s_i = r_i^* - \beta_i \phi_i\left(r_i^*|c\right) + v_i \tag{7}$$

Therefore, the equilibrium strategy $\rho_i : \mathcal{S}_i\left(c\right) \to \left[\underline{r}, \bar{r}\right]$ is invertible, hence strictly increasing, in the set $\mathcal{S}_i^I\left(c\right) = \left\{s_i \in \mathcal{S}_i\left(c\right) : \rho_i(s_i) > \underline{r}\right\}$. The points were $\rho_i(s_i) = \underline{r}$ are excluded from $\mathcal{S}_i^I\left(c\right)$ because the lower bound $\underline{r}$ might be binding for some s_i.

Proposition 1. *For each bank i, the equilibrium strategy $\rho_i : \mathcal{S}_i\left(c\right) \to \left[\underline{r}, \bar{r}\right]$ is strictly increasing at all $s_i \in \mathcal{S}_i^I\left(c\right)$*

Henceforth, I refer to Equation (7) as the *inverse equilibrium strategy* of bank i. Such inversion is crucial for proving some of the properties of the Bayesian Nash Equilibrium. Besides, it also provides the key for the identification of the model, following a strategy similar to Guerre et al. (2000), Li et al. (2000), Li et al. (2002), and all the subsequent literature in structural estimation of auctions, as will be shown below in the section on identification.

4 Identification Strategy

In this section, I present identification results for a fixed period p. Hence, I drop the index p but keep t to denote a day in the sample. Ideally, we would want the model in Section 3 to identify the distributions of the borrowing costs (at least conditional on c) and, if possible, for every bank i and every day t, the specific borrowing cost $s_{i,t}$. If we could recover all these costs, we would be able to compute the daily (trimmed) average borrowing cost of the N banks in the panel, which is precisely the value that the reference rate would have under truthful reporting (the unmanipulated Libor). Unfortunately, this cannot be done using data on quotes alone.[13] However, I conclude this section by showing that a lower bound on the extent of misreporting is identified. Furthermore, in Section 5, I propose an estimate of a lower bound on the common cost c_t and the unmanipulated Libor, which is quite useful given the evidence suggesting that Libor was manipulated downward during the period considered. In fact, as shown in Section 6, the estimated lower bound is mostly above the published rate during this period, which is consistent with previous evidence.

Let us start by considering a simple case where c_t is constant, and v_i and β_i are observed. In such a case, banks would be playing the same static game at each t, and for each bank i its observed quote $r_{i,t}$ would be drawn from the same unconditional distribution every day t. Hence, these distributions could be estimated directly from the observed quotes. Moreover, we could also estimate the probabilities $\phi_i\left(r_{i,t}|\bar{c}\right)$, from the estimated quote distributions. Then, we could recover the private costs of each bank, at each t, from the inverse equilibrium strategies:

$$s_{i,t} = r_{i,t} - \beta_i \phi_i\left(r_{i,t}|\bar{c}\right) + v_i \tag{8}$$

4.1 Unobserved Game Heterogeneity

More realistically,c_t changes every day and it is unobservable to the econometrician. Since it is common knowledge for the banks, it changes the game that banks are playing each day, hence the observed quotes follow different distributions each day. To control for this unobserved heterogeneity I follow a strategy similar to Krasnokutskaya (2011).

Following an argument analogous to that in Haile et al. (2003) and Krasnokutskaya (2011), Proposition 2 states that the equilibrium strategies (when thought of as functions of both the purely private costs $\pi_{i,t}$ and the common cost c_t), are additively separable into a common component and a private quote component.

[13] To better understand the challenges to the identification strategy, let's consider the first price sealed bid auction with independent private values analyzed in Guerre et al. (2000), where the private values of all bidders at each auction are indeed identified from the set of bids submitted by all bidders. As opposed to their case, here the inverse equilibrium strategies (7) depend on parameters (β_i and v_i), that are unknown to the econometrician. Moreover, the costs are not independent across banks, but only conditionally independent given a common cost c, that is common knowledge for the banks, but unobservable to the econometrician. Our setup is closer to that of Campo (2012), who admits heterogeneous risk aversion, and Krasnokutskaya (2011), who allows for unobserved auction heterogeneity. In the former, the risk aversion coefficient of each bidder is identified by imposing strong conditions on the distributions of their private values (they are all assumed to have the same marginal distributions). In the latter, the distributions of the private and the common cost components (in a procurement auction) are identified, but the specific realizations of these costs that correspond to the observed bids, are not.

Proposition 2. *Let $\rho_j(\cdot;0)$ be the equilibrium strategies of all banks $j \in \mathcal{N}$, when $c_t = 0$. If the strategy of each bank $j \neq i$, when $c_t \neq 0$, is*

$$\rho_j(\pi_{j,t}; c_t) = c_t + \rho_j(\pi_{j,t}; 0) \tag{9}$$

then i's best response is also $\rho_i(\pi_{i,t}; c_t) = c_t + \rho_i(\pi_{i,t}; 0)$.

The proof of Proposition 2 is left for Appendix A. It consists of directly verifying that $c_t + \rho_i(\pi_{i,t}; 0)$ maximizes i's expected payoff when all other players are using the strategies in (9). An immediate corollary of Proposition 2 is that, if $(\rho_1^0, ..., \rho_N^0)$ is a BNE of the game with $c_t = 0$, then $(\rho_1, ..., \rho_N)$, with $\rho_i = \rho_i^0 + c_t$ for all $i \in \mathcal{N}$, is a BNE of the game with $c_t \neq 0$.[14]

Let $\Pi_{i,t}$ be the private cost component of bank i at time t, and $Q_{i,t} = \rho_i(\Pi_{i,t}; 0)$ be the normalized equilibrium quote that bank i would submit, when its purely private cost is $\Pi_{i,t}$ and $C_t = 0$ (assuming these are random variables for the econometrician). It follows that, for all $i \in \mathcal{N}$, $R_{i,t} = C_t + Q_{i,t}$. In particular, for any two banks $i_1, i_2 \in \mathcal{N}$,

$$\begin{aligned} R_{i_1,t} &= C_t + Q_{i_1,t} \\ R_{i_2,t} &= C_t + Q_{i_2,t} \end{aligned} \tag{10}$$

Notice that $Q_{i_1,t}$ and $Q_{i_2,t}$ are functions of $\Pi_{i_1,t}$ and $\Pi_{i_2,t}$, respectively, hence they are independent. Moreover, from the point of view of the econometrician, C_t is also independent of $Q_{i_1,t}$ and $Q_{i_2,t}$. The goal now is to recover the unknown distributions of C_t, $Q_{i_1,t}$ and $Q_{i_2,t}$ from the observable distributions of $R_{i_1,t}$ and $R_{i_2,t}$. Under the additional assumption that C_t is independently drawn from the same distribution at all t, this is the same deconvolution problem introduced by Li and Vuong (1998), in the context of measurement error with multiple indicators, and later studied by Li et al. (2000) and Krasnokutskaya (2011) for the estimation of procurement auctions with a common cost component. All of them base their results on Kotlarski's Lemma (Kotlarski, 1966).

An additional difficulty here is that C_t could be drawn from a different distribution on each t and these draws may not be independent across time, hence the assumptions of Kotlarski's Lemma would no longer hold.[15] When C_t is not i.i.d. it is necessary to assume further that at least three quotes and the identities of the banks submitting them are observable.[16] Accordingly, I assume that the joint distribution of pairs of the form $(R_{i,t} - R_{j,t}, R_{k,t} - R_{j,t})$ is given.

[14] A critical assumption when estimating static games is that at each repetition of the game, the observable set of actions chosen by the players correspond to the same equilibrium (conditional on the unobserved heterogeneity). Of course, this assumption would be trivially satisfied if the equilibrium $(\rho_1^0, ..., \rho_N^0)$ was unique. If instead there were multiple equilibria for a given c, the identification strategy would fail because the same distribution of borrowing costs would result in a different distribution of quotes under different equilibria. The observed distribution of quotes would then be a mixture of these different distributions from different equilibria. To address this concern, I conducted simulations where I computed the BNE of the game numerically, for different regions of the parameters space. Reassuringly, the results suggest that the equilibrium is in fact unique. All simulations converge to the same equilibrium strategies, regardless of the starting point. A detailed description is included in Online Appendix B.

[15] Li et al. (2000) and Krasnokutskaya (2011) provide identification results for the case with i.i.d. C_t, which require that at least two bids and the identities of the corresponding bidders are observable.

[16] The data strongly suggests that the unobserved common cost component C_t is highly persistent and may even follow a non-stationary process, as shown in Online Appendix B. To the best of my knowledge, the particular case

Assumption 5. *For all $i, j, k \in \mathcal{N}$,the joint distribution of $(R_{i,t} - R_{j,t}, R_{k,t} - R_{j,t})$ is known. Here $R_{i,t}$, $R_{j,t}$ and $R_{k,t}$ are random variables denoting optimal quotes.*

Given three different bidders, i_1, i_2 and i_3, Equation (10) implies that

$$\begin{aligned} R_{i_1,t} - R_{i_2,t} &= -Q_{i_2,t} + Q_{i_1,t} \\ R_{i_3,t} - R_{i_2,t} &= -Q_{i_2,t} + Q_{i_3,t} \end{aligned}$$

where $Q_{i_1,t}$, $Q_{i_2,t}$ and $Q_{i_3,t}$ are mutually independent, and $Q_{i,t}$ is i.i.d. for each i, thus these two equations satisfy the assumptions of Kotlarski's Lemma. Therefore, Assumption 5 implies Proposition 3 below (the details of the proof are left for Appendix A).

Proposition 3. *If the characteristic functions of the normalized quotes $Q_{i_1,t}$, $Q_{i_2,t}$ and $Q_{i_3,t}$ are non-vanishing everywhere, their distributions are identified, up to an additive constant ($-E[Q_{i_1,t}]$), from the joint distribution of $(R_{i_1,t} - R_{i_2,t}, R_{i_3,t} - R_{i_2,t})$.*

It is worth noticing that the constant term $E[Q_{i_1,t}]$ in Proposition 3 is not identified. Hence, we only achieve identification of the distributions of the normalized quotes up to an additive constant. However, the distributions of $Q_{i_1,t} - E[Q_{i_1,t}]$, $Q_{i_2,t} - E[Q_{i_1,t}]$ and $Q_{i_3,t} - E[Q_{i_1,t}]$ are all identified. Let $\tilde{Q}_{i,t} = Q_{i,t} - E[Q_{i_1,t}]$, Proposition 2 implies $\tilde{Q}_{i,t}$ is the optimal report of bank i in a game with $c_t = -E[Q_{i_1,t}]$.

Proposition 3 is crucial for the identification of the model, because it removes the unobserved heterogeneity from the game. More importantly, it implies the distribution of the normalized quotes $\tilde{Q}_{i,t}$ are identified, even though the constant $-E[Q_{i_1}]$ remains unidentified.

4.2 Identified Parameters

Proposition 3 allows us to fix the common cost $c = -E[Q_{i_1}]$. I now establish conditions for the identification of β_i, v_i and the distribution of $\Pi_{i,t}$, for all $i \in \mathcal{N}$.

First, I notice that the probability $\phi_i(\tilde{q}_{i,t}| - E[Q_{i_1,t}])$ is identified for each value $\tilde{q}_{i,t}$ in the support of $\tilde{Q}_{i,t}$. Proposition 3 implies that the distribution of $\tilde{Q}_{j,t}$ is identified for all $j \in \mathcal{N}$ and it does not vary with t. It follows that,

$$\phi_i(\tilde{q}_{i,t}| - E[Q_{i_1,t}]) = P_{\tilde{Q}_{-i}}\left\{\tilde{Q}^{(\underline{n})} \leq \tilde{q}_{i,t} \leq \tilde{Q}^{(\bar{n})}\right\} \tag{11}$$

where $\tilde{Q}^{(\underline{n})}$ and $\tilde{Q}^{(\bar{n})}$ are order statistics of the vector of all normalized quotes. Therefore, these probabilities are identified for all $i \in \mathcal{N}$ and each realization $\tilde{q}_{i,t}$ of $\tilde{Q}_{i,t}$, because they are functions of identified distributions.

Moreover, $\phi_i(r_{i,t}|c_t) = \phi_i(\tilde{q}_{i,t}| - E[q_{i_1}])$, hence, despite the notation, identification of these probabilities does not require knowledge of c_t. Intuitively, since c_t is common knowledge for the

with non-i.i.d. heterogeneity has not been considered before in the empirical auctions literature, although it is just a natural extension of the model with i.i.d. heterogeneity.

banks and reports are additively separable (Proposition 2), the probability that i assigns to the event $r^{(\underline{n})} \leq r_i \leq r^{(\bar{n})}$ is equivalent to the probability that $q^{(\underline{n})} \leq q_i \leq q^{(\bar{n})}$. Formally,

$$\begin{aligned}
\phi_i\left(r_{i,t}|c_t\right) &= P_i\left\{r_t^{(\underline{n})} \leq r_{i,t} \leq r_t^{(\bar{n})}|c_t\right\} \\
&= P_i\left\{r_t^{(\underline{n})} - c_t \leq r_{i,t} - c_t \leq r_t^{(\bar{n})} - c_t\right\} \\
&= P_i\left\{q_t^{(\underline{n})} \leq q_{i,t} \leq q_t^{(\bar{n})}\right\} \\
&= P_i\left\{q_t^{(\underline{n})} - E\left[q_{i_1}\right] \leq q_{i,t} - E\left[q_{i_1}\right] \leq q_t^{(\bar{n})} - E\left[q_{i_1}\right]\right\} \\
&= \phi_i\left(q_{i,t} - E\left[q_{i_1}\right] | - E\left[q_{i_1}\right]\right)
\end{aligned} \tag{12}$$

Given the results in 11 and 12, I will used the simplified notation $\phi_i\left(\tilde{q}_{i,t}\right)$ to denote the probability $\phi_i\left(\tilde{q}_{i,t}| - E\left[Q_{i_1,t}\right]\right)$.

Let us now derive an expression for β_i in terms of the distribution of $\tilde{Q}_{i,t}$ and the probability $\phi_i\left(\tilde{q}_{i,t}\right)$. The first order condition (7) of the game with $c = -E\left[Q_{i_1,t}\right]$ can be written as:

$$-c + \pi_{i,t} = \tilde{q}_{i,t} - \beta_i \phi_i\left(\tilde{q}_{i,t}\right) + v_i \tag{13}$$

for all $i \in \mathcal{N}$. Moreover, taking expectations,

$$-c + E\left[\Pi_{i,t}\right] = E\left[\tilde{Q}_{i,t}\right] - \beta_i E_{\tilde{Q}_i}\left[\phi_i\left(\tilde{Q}_{i,t}\right)\right] + v_i \tag{14}$$

and subtracting expected values from 13 we get,

$$\Pi_{i,t} \quad -E\left[\Pi_{i,t}\right] = \quad \tilde{Q}_{i,t} - E\left[\tilde{Q}_{i,t}\right] - \beta_i\left(\phi_i\left(\tilde{Q}_{i,t}\right) - E_{\tilde{Q}_i}\left[\phi_i\left(\tilde{Q}_{i,t}\right)\right]\right) \tag{15}$$

The distribution of $\tilde{Q}_{i,t}$ is identified and does not vary with t, thus the means $E\left[\tilde{Q}_{i,t}\right]$ and $E_{\tilde{Q}_i}\left[\phi_i\left(\tilde{Q}_{i,t}\right)\right]$ are identified as well. Equation (15) is the key for the identification of β_i (and the distribution of $\Pi_{i,t} - E\left[\Pi_{i,t}\right]$). However, an additional restriction must be imposed on F_i to achieve identification.

Assumption 6. *For all $i \in \mathcal{N}$ $Med\left(\Pi_{i,t}\right) = E\left[\Pi_{i,t}\right]$, where $Med\left(\Pi_{i,t}\right)$ denotes the median of $\Pi_{i,t}$.*

Assumption 6 states that, for any given bank, it is equally likely that it receives a positive or a negative idiosyncratic shock. As a restriction on the distributions of $\Pi_{i,t}$, this assumption is implied by, but weaker than, symmetry around the mean. Under such assumption, the parameter β_i and the distribution of $\Pi_{i,t}$ are identified, for all i, as stated in the next two propositions.

Assumption 7. *In the game with $c_t = -E\left[Q_{i_1,t}\right]$, the optimal quote of bank i when $\pi_{i,t} = E\left[\Pi_{i,t}\right]$ is an interior solution to (3), for all $i \in \mathcal{N}$. Briefly, $0 \in \mathcal{S}_i^I\left(-E\left[Q_{i_1,t}\right]\right)$,for all $i \in \mathcal{N}$.*

Assumption 7 requires that the lower bound $\underline{r}$ in the set of admissible reported rates is not binding when the private cost is equal to its mean. In the model, we can make $\underline{r}$ low enough (even negative) to ensure it is never binding. More realistically, when the model is matched to the data, a reasonable bound is $\underline{r} = 0$. Reassuringly, Libor rates have never hit the zero lower bound so this restriction is innocuous.

Proposition 4. *Under Assumptions 2-7, the preference parameter β_i is identified, for all $i \in \mathcal{N}$, from the distributions of the normalized quotes. Moreover,*

$$\beta_i = \frac{Med\left(\tilde{Q}_{i,t}\right) - E\left[\tilde{Q}_{i,t}\right]}{\phi_i\left(Med\left(\tilde{Q}_{i,t}\right)\right) - E_{\tilde{Q}_i}\left[\phi_i\left(\tilde{Q}_{i,t}\right)\right]} \tag{16}$$

where $\tilde{Q}_{i,t}$ is a random variable that denotes the quote of bank i in the game with $c_t = -E\left[Q_{i_1,t}\right]$. The distribution of $\Pi_{i,t} - E\left[\Pi_{i,t}\right]$ is identified as well.

The proof of (16) is left for Appendix A. It relies on the fact that Equation (15) is a translation of the inverse equilibrium strategy of bank i in the game with $c_t = -E\left[Q_{i_1,t}\right]$, hence it is satisfied by any interior optimal normalized quote. Moreover, since the optimal strategy is strictly increasing in the interior of the support of $\Pi_{i,t}$, when the private cost is equal to $\text{Med}\left(\Pi_{i,t}\right)$ the optimal quote is $\text{Med}\left(\tilde{Q}_{i,t}\right)$. Identification of the distribution of $\Pi_{i,t} - E\left[\Pi_{i,t}\right]$ follows directly from Proposition 4 and Equation (15), since all the distributions and the constants on the right hand side of this equation are identified.

Additional assumptions are required to recover all other model primitives. Notice that $-E\left[Q_{i_1,t}\right]$, $E\left[\Pi_{i,t}\right]$ and v_i are all unknown additive constants in Equation (14), hence they cannot be separately identified from reports alone. I thus add the next assumption.

Assumption 8. *For each bank i, the mean of its purely private cost is zero, that is, $E\left[\Pi_{i,t}\right] = 0$.*

Assumption 8 rules out any first order heterogeneity in the distributions of banks' borrowing costs.However, such a model still allows for persistent differences in risk, since the distributions might have different variances. It might be the case that some banks in the USD Libor panel face persistently higher borrowing costs than others. To address this concern, I report in Online Appendix B the results of an estimation of a version of the model where such differences are allowed (a model where Assumption 8 does not hold). However, that model restricts the signaling parameter v_i to be the same for all banks and it requires additional data sources for identification. In Section 6 I compare the results of the estimation of these two models.[17]

Proposition 5. *Under Assumptions 2-8, the distribution of $\Pi_{i,t}$, F_i, and the deviation of the private signaling parameters from their mean, $v_i - \bar{v}$, are identified for For all $i \in \mathcal{N}$. Moreover, for all $i \in \mathcal{N}$, a lower bound on its private signaling parameter v_i is also identified.*

[17] Assumption 8 and the normalization $c = 0$ in Proposition 2 imply that I am allowing for negative borrowing costs, which seems unrealistic. However, such normalization is completely arbitrary and I chose zero just for mathematical convenience. I could have chosen any other fixed value $\bar{c}$ such that $\bar{c} + \underline{\pi} > 0$ to avoid negative borrowing costs, and all the identification results would remain the same.

Identification of F_i follows straightforward from Proposition 4 and Assumption 8. Equation (14) and Assumption 8 imply that $v_i - E\left[Q_{i_1,t}\right]$ is also identified for all $i \in \mathcal{N}$ (but $-E\left[Q_{i_1,t}\right]$ remains unknown). Moreover, for each $i \in \mathcal{N}$,

$$v_i - E\left[Q_{i_1,t}\right] - \min_{i\in\mathcal{N}} \left\{v_i - E\left[Q_{i_1,t}\right]\right\} = v_i - \min_{i\in\mathcal{N}} \left\{v_i\right\} \leq v_i \tag{17}$$

therefore $v_i - E\left[Q_{i_1,t}\right] - \min_{i\in\mathcal{N}} \left\{v_i - E\left[Q_{i_1,t}\right]\right\}$ identifies a lower bound on v_i.

4.3 Motives for misreporting and the distribution of normalized quotes

To illustrate the ability of the model to separately identify signaling and exposure to Libor as distinct motives for misreporting, I simulate the optimal response of Bank 1 to the equilibrium strategies of all other banks, under hypothetical scenarios with different values for the parameters v_1 and β_1, and a given distribution of private costs that satisfies Assumptions 2, 3 and 6 - 8.

First, I set $v_1 = 30$ bp, $\beta_1 = 10$bp, and $v_j = \beta_j = 0$ for all $j \neq 1$, and compute the BNE of the game described in Section 3. I provide a detailed description of the equilibrium simulations in Online Appendix B.

Let us remember from 6, that the best response of any bank i, ρ_i, is implicitly defined by $\rho_i\left(s_i\right) = s_i + \beta_i \phi_i\left(\rho_i\left(s_i\right)|c\right) - v_i$. Figure 2 shows the two optimal responses of Bank 1 to the equilibrium strategies of all other banks, under two different sets of parameters. The right panel shows the equilibrium strategy with $v_1 = 30$ and $\beta_i = 10$. For comparison, the left panel shows the best response of the same bank with $v_1 = 0$ and $\beta_1 = 10$, keeping constant all other parameters and the distributions of borrowing costs. When $v_1 = 0$, the only incentive for misreporting is the exposure to Libor. In such a case, the bank optimally chooses to report its costs truthfully when the probability ϕ_i that its report is included in the benchmark rate computation is either zero or one. When its costs are closer to the mean (zero), the bank assigns a higher probability to this event and thus has higher incentives to overstate its costs and benefits from its positive exposure to Libor. The effect of adding signaling as a further motive for misreporting, while keeping $\beta_i = 10$ constant, is just a vertical translation of the equilibrium strategy (by exactly-30 bp). The benefit the bank derives from signaling is the same regardless of its true borrowing cost.

Similarly, the left panel of Figure 3, shows the difference between the distributions of the costs and optimal reports, when $\beta_1 = 0$ and $v_1 = 30$. In this case, the distribution of reports is just a horizontal translation of the distribution of costs. When $\beta_1 = 0$, the optimal strategy simplifies to $\rho_1\left(s_1\right) = s_1 - v_1$; hence, the extent of misreporting, exclusively due to signaling, is the same across the entire support of the distribution of its private costs. In contrast, when $\beta_1 \neq 0$, the size of misreporting from exposure to Libor depends on ϕ_i, hence it is a non-constant function of the bank's private cost. The right panel of Figure 3 shows the resulting distribution of reported rates when $v_1 = 30$ and $\beta_1 = 10$. If the exposure is sufficiently large to induce misreporting, it changes the shape of the distribution of the normalized quotes when compared to the distribution of the private costs. In particular, the former is no longer symmetric; hence, its mean and median no

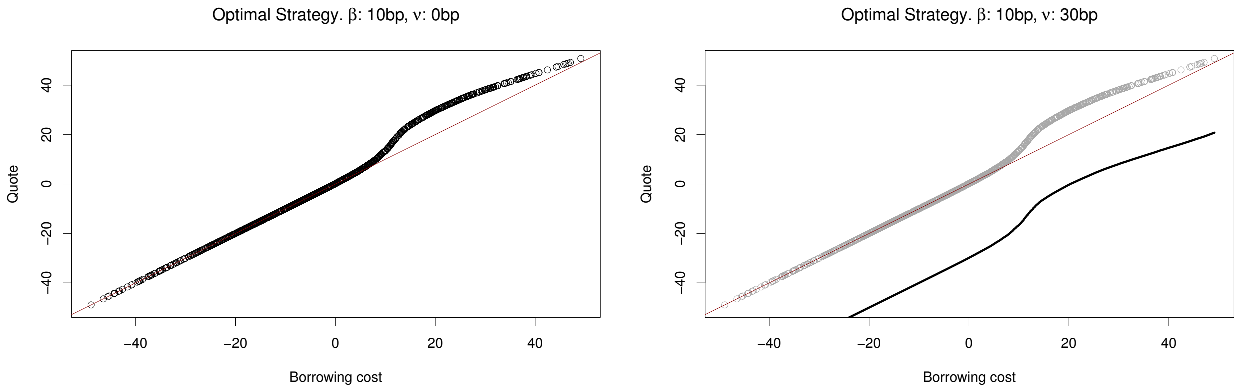

Figure 2: Best responses to equilibrium strategies with different incentives for misreporting

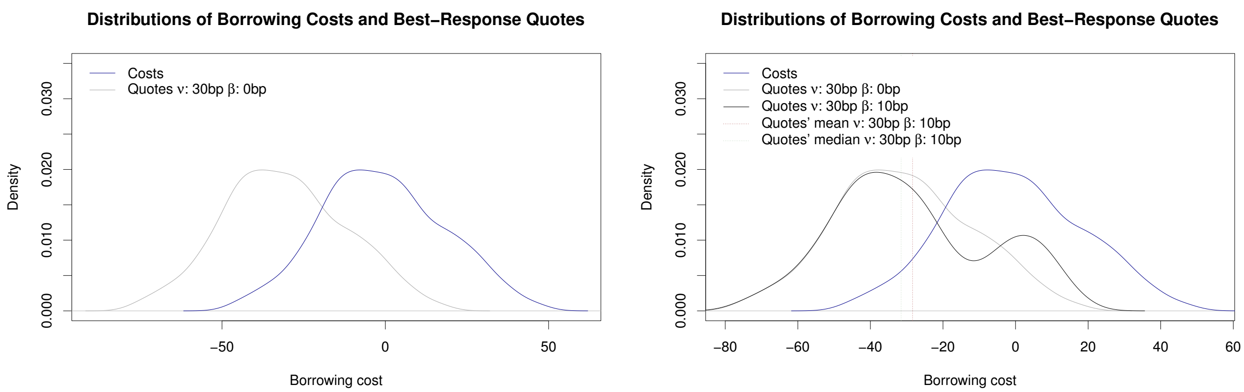

Figure 3: Equilibrium distributions of quotes under different incentives for misreporting

longer coincide.

In a nut shell, the model attributes any observed difference between the mean and the median of the distribution of normalized quotes to β_1 (given Assumption 6) and identifies (a lower bound for) v_1 from any change in location (translation) that remains unexplained after controlling for β_1 (given Assumption 8). This highlights the importance of Assumption 6 for the identification of the signaling parameters. If the distributions of idiosyncratic costs all have the same mean, then the model attributes the unexplained differences in the average costs reported by banks to signaling.

4.4 Lower bound for the extent of misreporting and manipulation

Besides the signaling and exposure parameters in the payoff function of the banks, under Assumptions 2 - 8, average measures of the extent of misreporting for each bank and the overall Libor manipulation are also identified. Notice that

$$E\left[s_{i,t} - r_{i,t}\right] = v_i - \beta_i E\left[\phi_i\left(r_{i,t}|c_t\right)\right] \tag{18}$$

and all terms on the right hand side of Equation (18) are identified, except for v_i for which only a lower bound is identified. Thus we can also identify a lower bound on each banks' average misreporting. Moreover, the model also identifies a lower bound on the mean spread between average borrowing costs and average Libor quotes, $E\left[\frac{1}{N}\sum_{i\in\mathcal{N}}(s_{it}-r_{i,t})\right]$, that provides a conservative measure of the extent of Libor manipulation.

Notice that I have not provided any arguments showing that the daily realizations of the common cost component c_t are identified. In fact, the identification strategy described in this section consists of removing the game heterogeneity due to c_t from each instance of the game, so that they can be considered independent occurrences of the same game. This allows us to identify the distributions of the normalized quotes; however, the corresponding normalization implies that the results are invariant to additive constants that affedt all costs equally. Therefore, daily measures of the unmanipulated Libor are not identified by these procedure. [18]

5 Estimation

The model can be estimated using a procedure based on the deconvolution result presented in Section 4. Very similar estimation algorithms are proposed in Li and Vuong (1998), Li et al. (2000) and Krasnokutskaya (2011). Li et al. (2000) establish conditions under which the resulting estimators are uniformly consistent. However, these estimators are known to have very low convergence rates, as shown by Li and Vuong (1998), even under strong assumptions about the smoothness of the distributions involved.[19] Broadly speaking, the moral is that for short samples (say, less than 500 observations) the density estimators would have non-negligible noise and would be highly sensitive to the choice of a particular smoothing method.[20] As an alternative, I propose an estimation method that has a noise term that does not vanish asymptotically as the number of days in the sample increases (unless the number of banks also grows large), but still has low variance relative to the variance of the quotes, as I show below.[21]

The estimation procedure in this section must be applied separately for each period p; hence, I drop the corresponding index. Roughly, the main idea is to use the observable deviations of the individual quotes from their intraday average as (moderately) noisy measures of normalized quotes. From Section 4, under maintained assumptions, the quotes are additively separable into a common component and a purely private component. That is, $r_{i,t} = c_t + q_{i,t}$, for all $i \in \mathcal{N}$ and all t.

[18] In Section 5, I propose a (noisy) estimator of a lower bound of the daily common cost component c_t. In Section 6, I show that the estimated lower bound always lays above the reported Libor, and mostly in-between two other measures of average borrowing costs already described in Section 2.

[19] See Ch. 5 in Horowitz (2009) for a general description of the deconvolution estimator and a discussion about the slow rate of convergence of the corresponding density estimator.

[20] The empirical characteristic function must be smoothed to guarantee the existence of the integral in the inverse Fourier transform, which requires the econometrician to make some choices regarding smoothing functions and parameters

[21] I have performed several Monte Carlo simulations in which the method I propose in Section 5.1 outperforms the estimation based on deconvolution, even for samples larger than $T = 500$. The results of these simulations are reported in Online Appendix B.

Therefore, $\bar{r}_t^w = c_t + \bar{q}_t^w$, where $\bar{r}_t^w = \sum_{i\in\mathcal{N}} w_i r_{i,t}$, with $w_i > 0$ for all i and $\sum_{i\in\mathcal{N}} w_i = 1$, and $\bar{q}_t^w$ is analogously defined. Let $\xi_t^w = \bar{q}_t^w - E\left[\bar{q}^w\right]$, it follows that

$$r_{i,t} - \bar{r}_t^w = q_{i,t} - E\left[\bar{q}^w\right] - \xi_t^w \tag{19}$$

where, by construction, $E\left[\xi_t\right] = 0$ and $Var\left[\xi_t^w\right] = Var\left[\bar{q}_t^w\right] = \sum_{i\in\mathcal{N}} w_i^2 Var\left[q_{i,t}\right]$.[22]

5.1 Estimation Procedure

By subtracting the intraday mean from the observed quotes, we remove the unobserved heterogeneity almost entirely, except for the remaining mean-zero error term ξ_t^w. It is easy to show that the weights that minimize the variance of ξ_t^w are the inverse of the variance of the normalized quotes, that is, $w_i = Var\left[q_{i,t}\right]^{-1}\left(\sum_{i=1}^N Var\left[q_{i,t}\right]^{-1}\right)^{-1}$. In Appendix A, I show how to obtain a consistent estimate of $Var\left[q_{i,t}\right]$, denoted $\hat{\sigma}_{q_i}^2$. I use these estimates to compute the weighted average $\bar{r}_t^w$. Moreover, $\hat{Var}\left[\xi_t^w\right] = \left(\sum_{i\in\mathcal{N}} \frac{1}{\hat{\sigma}_{q_i}^2}\right)^{-1}$ provides an estimate of $Var\left[\xi_t^w\right]$.

For any bank i, the observed $r_{i,t} - \bar{r}_t^w$ are independent draws from the distribution of the normalized quote $q_i - E\left[\bar{q}^w\right]$ (plus noise $-\xi_t^w$). Thus, the entire sample $\{r_{i,t} - \bar{r}_t^w\}_{t=1}^T$ can be used to estimate the distribution of $q_i - E\left[\bar{q}^w\right]$, for each $i \in \mathcal{N}$. Section 4 shows that the distributions of $q_i - E\left[q_{i_1}\right]$ are identified. Equivalently, the distributions of $q_i - E\left[\bar{q}^w\right]$ are identified, since

$$q_i - E\left[\bar{q}^w\right] = q_i - E\left[q_{i_1}\right] - \sum_{i\in\mathcal{N}} w_i E\left[q_i - E\left[q_{i_1}\right]\right]$$

All the results concerning (the distributions of) $q_i - E\left[q_{i_1}\right]$ also hold for $q_i - E\left[\bar{q}^w\right]$. Note that in Section 4.2 we could have considered the hypothetical game with $c = -E\left[\bar{q}^w\right]$ (instead of $c = -E\left[q_{i_1}\right]$) and all the results would hold.

Besides having a lower variance in small samples, the main advantage of this alternative estimation procedure compared to the one based on deconvolution is that it allows us to recover (with low-variance, mean-zero error) the actual realizations of the normalized quotes for each t, and not just their distributions. That is, suppose that on day t, bank i observes the cost $s_{i,t} = c_t + \pi_{i,t}$, for a specific realization of $\pi_{i,t}$, and correspondingly submits a quote $r_{i,t}$. Then, $r_{i,t} - \bar{r}_t^w$ is a (noisy) estimate of the quote that bank i would have submitted in a counterfactual game with $c_t = -E\left[\bar{q}^w\right]$, if it had faced the cost $s_{i,t} = -E\left[\bar{q}^w\right] + \pi_{i,t}$ (for the same realization of $\pi_{i,t}$). Thus, for each t, $(r_{1,t} - \bar{r}_t^w, ..., r_{N,t} - \bar{r}_t^w)$ is an estimate of the vector of realized normalized quotes, corresponding to the actual purely private costs observed by banks $(\pi_{1,t}, ..., \pi_{N,t})$.

Let $\tilde{q}_{i,t} = r_{i,t} - \bar{r}_t^w$, then it follows from the previous argument that we have the whole sample $\{(\tilde{q}_{1,t}, ..., \tilde{q}_{N,t})\}_{t=1}^T$ of realized normalized quotes available for the estimation of the model. Equation

[22] Even though $E\left[\xi_t\right] = 0$, it should be noted that ξ_t cannot be interpreted as classical measurement error, since it is correlated with $q_{i,t}$. Moreover, ξ_t does converge in probability to zero as the number of banks increases.

(12) implies that a natural estimator of $\phi_i\left(r_{i,t}|c_t\right)$ would be:

$$\hat{\phi}_i\left(r_{i,t}|c_t\right) = \frac{1}{T}\sum_{\tau=1}^{T}\mathbf{1}\left(\tilde{q}_\tau^{(\underline{n})} \leq \tilde{q}_{i,t} \leq \tilde{q}_\tau^{(\bar{n})}\right) \tag{20}$$

where $\tilde{q}_\tau^{(\underline{n})}$ and $\tilde{q}_\tau^{(\bar{n})}$ are order statistics of the vector of all (normalized) quotes.

To increase the efficiency of $\hat{\phi}_i\left(r_{i,t}|c_t\right)$, I use a resampling method based on Hortaçsu (2000) and Hortaçsu and McAdams (2010). The procedure works as follows. Fix any bank i, and for each other bank $j \neq i$, take T_s draws with replacement from the sample $\{\tilde{q}_{j,t}\}_{t=1}^{T}$ of normalized quotes. Let $q_{j,\tau}^{s}$ denote the τ-th of such draws. Then, build T_s possible scenarios for bank i, that is T_s vectors of normalized quotes $\left\{q_{-i,\tau}^{s}\right\}_{\tau=1}^{T_s}$ that the other banks could have submitted, given the distribution of their normalized quotes which is known by bank i in equilibrium. Finally, for each $r_{i,t}$ estimate the probability $\phi_i\left(r_{i,t}|c_t\right)$ using (20) after replacing $\tilde{q}_\tau^{(\underline{n})}$, $\tilde{q}_\tau^{(\underline{n})}$ and T by $\tilde{q}_\tau^{s(\underline{n})}$, $\tilde{q}_\tau^{s(\underline{n})}$ and T^s, respectively.

With an estimate of $\phi_i\left(r_{i,t}|c_t\right)$ at hand, the preference parameter β_i can be estimated using the sample counterpart of the expression for β_i in Proposition 4. Suppose, for now, that we have such an estimate $\hat{\beta}_i$, for all $i \in \mathcal{N}$. Then, the private cost component $\pi_{i,t}$ observed by each bank i at each date t in the sample can be estimated by

$$\hat{\pi}_{i,t} = \tilde{q}_{i,t} - \frac{1}{T}\sum_{t=1}^{T}\tilde{q}_{i,t} - \hat{\beta}_i\left(\hat{\phi}_i\left(\tilde{q}_{i,t}\right) - \frac{1}{T}\sum_{t=1}^{T}\hat{\phi}_i\left(\tilde{q}_{i,t}\right)\right) \tag{21}$$

where $\hat{\phi}_i\left(\tilde{q}_{i,t}\right) = \hat{\phi}_i\left(r_{i,t}|c_t\right)$ as in (20). This is the empirical counterpart of Equation (15).

Similarly, an estimate of $v_i - \bar{v}$, the signaling parameter of bank i, is derived from the empirical counterpart of Equation (13):

$$\hat{v}_i = -\frac{1}{T}\sum_{t=1}^{T}\left(\tilde{q}_{i,t} - \frac{1}{N}\sum_{i\in\mathcal{N}}\tilde{q}_{i,t}\right) - \hat{\beta}_i\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\phi}_i\left(\tilde{q}_{i,t}\right) - \frac{1}{N}\sum_{i\in\mathcal{N}}\hat{\beta}_i\hat{\phi}_i\left(\tilde{q}_{i,t}\right)\right)$$

Moreover, lower bounds on each v_i are obtained by setting $\underline{\hat{v}}_i = \hat{v}_i - \min\left\{\hat{v}_i : i \in \mathcal{N}\right\}$.

We can then estimate lower bounds for all common costs $\{c_t\}_{t=0}^{T}$, since $c_t = r_{i,t} - \beta_i\phi_i\left(r_{i,t}|c_t\right) + v_i - \pi_{i,t}$ and we either observe, or can estimate, all the terms on the right hand side of this expression. An estimate of these lower bounds is given by

$$\bar{c}_t = \frac{1}{N}\sum_{i=1}^{N}\left(r_{i,t} - \hat{\beta}_i\hat{\phi}_i\left(\tilde{q}_{i,t}\right) + \underline{\hat{v}}_i - \hat{\pi}_{i,t}\right) \tag{22}$$

5.2 Estimation of β_i

The identification of β_i relies on the assumption that $\text{Med}\left(\pi_i\right) = 0$. Proposition 4 shows that under this assumption β_i can be explcitly expressed as a function of the distributions of the normalized quotes. We can derive a point estimator of β_i from this proposition, but before I derive a result of the model that is useful for estimating β_i because it places bounds on the values of β_i that are consistent with the optimality of the observed quotes $r_{i,t}$.

Assuming differentiability of the best response ρ_i,[23] for all $s_i \in \mathcal{S}_i(c)$, $\frac{d\rho_i(s_i)}{ds_i} = \left(1 - \beta_i \frac{d\phi_i(\rho_i(s_i)|c)}{dr_i}\right)^{-1}$. Since ρ_i is strictly increasing in $\mathcal{S}_i(c)$, $\frac{d\rho_i(s_i)}{ds_i} > 0$ and thus $1 - \beta_i \frac{d\phi_i(\rho_i(s_i)|c)}{dr_i} > 0$, for all $s_i \in \mathcal{S}_i(c)$. Therefore, for all optimal r_i, in particular, for any equilibrium quote r_i, $1 - \beta_i \frac{d\phi_i(r_i|c)}{dr_i} > 0$. Equivalently, for all r_i, $\beta_i < 1/\frac{d\phi_i(r_i|c)}{dr_i}$ if $\frac{d\phi_i(r_i|c)}{dr_i} > 0$ and $\beta_i > 1/\frac{d\phi_i(r_i|c)}{dr_i}$ if $\frac{d\phi_i(r_i|c)}{dr_i} < 0$.

It follows that,

$$1/\inf_{q_i \in Q_i^-}\left(\frac{d\phi_i(q_i)}{dq_i}\right) \leq \beta_i \leq 1/\sup_{q_i \in Q_i^+}\left(\frac{d\phi_i(q_i)}{dq_i}\right) \tag{23}$$

where $Q_i^- = \left\{q_i \in \left(\underline{q}_i, \bar{q}_i\right) : \frac{d\phi_i(q_i)}{dq_i} < 0\right\}$ and $Q_i^+ = \left\{q_i \in \left(\underline{q}_i, \bar{q}_i\right) : \frac{d\phi_i(q_i)}{dq_i} > 0\right\}$.

The inequalities in (23) place bounds on the values of β_i that rationalize the observed (normalized) quotes. Then, they hold for any β_i that satisfies (16) in Proposition 4. In a finite sample, however, the point estimate derived from (16) could lie outside these bounds. Therefore, to obtain estimates that are consistent with the assumption that all observed quotes correspond to the BNE of the game, I define an estimator of β_i that always lies within estimates of these bounds in (23). Let us denote these estimates of the lower and the upper bound $\hat{\beta}_i^{lb}$ and $\hat{\beta}_i^{ub}$, respectively. The idea is to find a value $\hat{\beta}_i$ in the interval $\left[\hat{\beta}_i^{lb}, \hat{\beta}_i^{ub}\right]$ that minimizes the squared median of the sample of private borrowing costs implied by the inverse equilibrium strategies and the observed distributions of the normalized quotes.

Formally, for any $b \in \left[\hat{\beta}_i^{lb}, \hat{\beta}_i^{ub}\right]$, let

$$\pi_{i,t}(b) = \tilde{q}_{i,t} - \frac{1}{T}\sum_{t=1}^{T} \tilde{q}_{i,t} - b\left(\hat{\phi}_i(\tilde{q}_{i,t}) - \frac{1}{T}\sum_{t=1}^{T}\hat{\phi}_i(\tilde{q}_{i,t})\right)$$

The estimator $\hat{\beta}_i$ is defined by

$$\hat{\beta}_i = \underset{b \in \left[\hat{\beta}_i^{lb}, \hat{\beta}_i^{ub}\right]}{\operatorname{argmin}} \left(\operatorname{Med}\left\{\pi_{i,t}(b) : t = 1, ..., T\right\}\right)^2 \tag{24}$$

which is a classical minimum distance estimator (see Newey and McFadden, 1994 for a general description of minimum distance estimators and its asymptotic properties). I do not derive here the asymptotic distribution of $\hat{\beta}_i$; instead, I use the bootstrap to estimate its variance and perform inference, as described in Online Appendix B.

To calculate $\hat{\beta}_i^{lb}$ and $\hat{\beta}_i^{ub}$, I first estimate the derivative $\frac{d\phi_i(\tilde{q}_{i,t})}{dq_i}$ for each estimated normalized quote $\tilde{q}_{i,t}$. Then I compute the empirical counterparts of the bounds in 23. It is possible, though, that the set $Q_i^-(Q_i^+)$ is empty for some banks. In such a case, the lower (upper) bound is just $-B$ (B), for an arbitrarily large B.

[23] I assume differentiability of the best response functions only for notational convenience, but bounds for β_i can also be derived from weaker assumptions using only one-sided derivatives.

6 Results

The model in Section 3 is estimated using only three-month USD Libor quotes for five different sample periods covering the financial crisis: (i) 09/03/2007 - 03/15/2008, (ii) 03/16/2008 - 09/14/2008, (iii) 09/15/2008 - 12/31/2008, (iv) 02/09/2009 - 12/31/2009, and (v) 01/01/2010 - 12/31/2010. (i) Starts the first trading day after BNP Paribas announced that it would suspend redemption in two subprime mortgage funds and ends with Bear Stearns failure. (iii) Goes from the day Lehman filed for bankruptcy until the end of 2008, when the BBA changed the composition of the USD Libor panel.[24] (iv) and (v) are somehow arbitrary periods of roughly one year each overlapping the manipulation scandal. The main criterion to choose these periods is that at the reference dates there is an apparent structural change in borrowing costs, or there are reasons to expect changes in banks' incentives to misreport (see Figure 1). Within each period, I estimate a lower bound for the unmanipulated Libor, the daily idiosyncratic cost shocks faced by each bank, and (bounds on) all parameters in the banks' payoff function. All parameters and distributions are held constant for each period. I restrict their duration to no more than one year to allow for potential variations in banks' incentives or in the distributions of their idiosyncratic borrowing costs. In principle, the model could be estimated separately with data from more but shorter periods to capture more frequent changes in those parameters, at the cost of decreasing the precision of the estimators.

6.1 Estimation of a counterfactual unmanipulated Libor

This section reports an estimate of the unmanipulated Libor, that is, the value that Libor would have had if banks had not misreported their borrowing costs. Since previous literature has found indications of downward manipulation, the lower bound is likely a conservative estimate of the extent of Libor manipulation. The left panel of Figure 4 shows that the published Libor is lower than the lower bound on the unmanipulated Libor on all days in my sample. In period (iii), the average Libor - unmanipulated Libor spread is -22 bp, and it reaches a maximum absolute deviation of 30 bp on 09/24/2008, only seven days after Lehman's bankruptcy. There is weaker evidence of manipulation in periods (i) and (ii), with an average spread of roughly -2 bp, but yet statistically lower than zero at a significance level of 0.05. Finally, the average spreads in periods (iii) and (iv) are -7 bp and -8 bp, respectively.

The right panel in Figure 4 compares the estimated unmanipulated Libor to the two other measures of average borrowing costs in the interbank market described in Section 2. On 85% of the days after 06/02/2008, when ICAP started publishing the NYFR, the estimated bound on the unmanipulated Libor lies between the other two measures and is consistently closer to the NYFR, except for the last period when all three measures yield similar results. Since the lower bound is estimated from banks' Libor quotes alone, this result validates the model based on testable implications, under the assumption that the NYFR and the EDR are more accurate measures of

[24] Since banks are ex-ante heterogeneous, a change in the Libor panel implies a new game structure. Therefore, I separately estimate the model before and after this change.

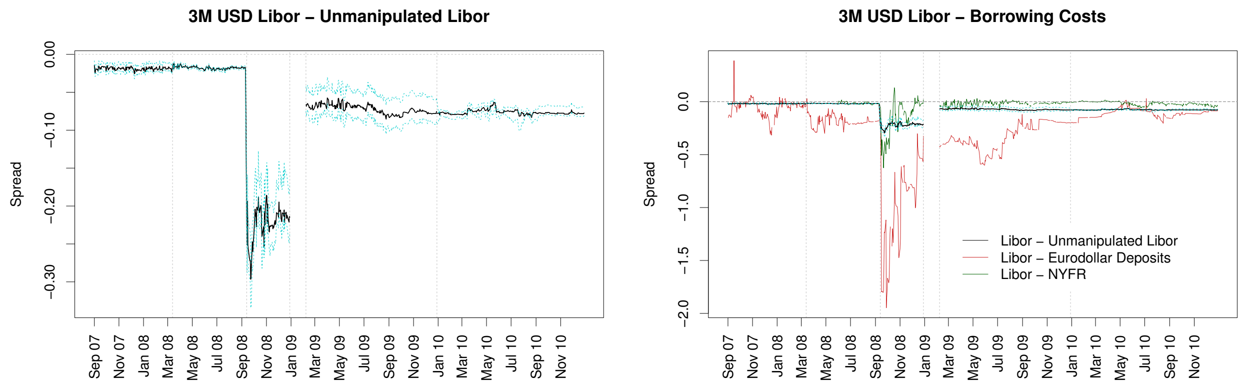

Figure 4: Spreads between Libor and the Estimated Common Borrowing Costs

The left panel shows the spread between the published Libor and the estimated unmanipulated Libor. The highlighted dates are 03/16/2008, when JPMorgan Chase acquired Bear Stearns for less than 7% of its market value two days before; 09/15/2008, when Lehman Brothers filed for bankruptcy; 12/31/2008, when the British Banking Authority changed the composition of the Libor panel; 02/09/2009; and 12/31/2009. Data on Libor quotes from 01/01/2009 to 02/09/2009 are not available. I computed 95% confidence intervals (light blue dotted lines) using the bootstrap. The right panel compares the estimated unmanipulated Libor to the two other measures of average borrowing costs described in Section 2 and Figure 1.

borrowing costs than the distorted Libor.

Although most empirical research on Libor manipulation has focused on the financial crisis when the signaling incentive was presumably stronger, related court documents contain evidence of manipulation attempts going back to 2005. To test whether these attempts succeeded in manipulating Libor, I extend the sample to include two more consecutive periods, from 09/03/2005 to 09/02/2006, and from 09/03/2006 to 09/02/2007. The estimated lower bound on the unmanipulated Libor provides very weak evidence, if any at all, of Libor manipulation during 09/03/2005 - 09/02/2007. Figure 5 shows the results. For this period, the average estimated spread between Libor and the lower bound on c_t is -0.6 bp. During this same period, the EDR is much closer to Libor than in any other period in the sample, and lies mostly above Libor. A likely explanation is that before the BNP Paribas announcement, there were no serious concerns about large financial institutions facing credit or liquidity risk; thus, banks did not have strong incentives to use Libor reports as signals of creditworthiness. Moreover, their individual attempts to benefit from trading positions may not have been sufficient to cause a substantial distortion in the benchmark rate.

6.2 Incentives for Misreporting

The size and direction of a bank's misreporting, that is, the difference between its quote and its borrowing cost, can be separated into two distinct components, one due to exposure to Libor and another exclusively attributable to signaling, as follows:

$$r_{i,t,p} - s_{i,t,p} = \underbrace{\frac{\alpha_{i,p}}{\tilde{N}}\phi_{i,p}\left(r_{i,t,p}|c_{t,p}\right)}_{\text{Exposure}} + \underbrace{\frac{v_{i,p}}{\tilde{N}}\phi_{i,p}\left(r_{i,t,p}|c_{t,p}\right) - v_{i,p}}_{\text{Signaling}} \tag{25}$$

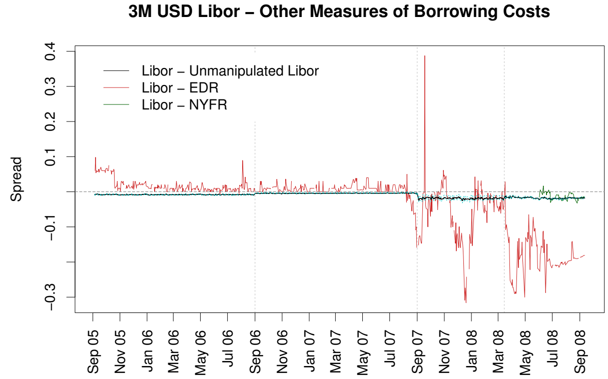

Figure 5: Spread between Libor and the Estimated Common Borrowing Costs Before the Crisis
This figure shows the spread between the published Libor and the estimated unmanipulated Libor from 09/03/2005 to 09/14/2008. It also includes the Libor - Eurodollar Deposits Rate (EDR) and Libor - New York Funding Rate (NYFR) spreads for comparison. The NYFR is only available from 06/02/2008. The highlighted dates are 09/03/2006, 09/03/2007 (the first trading day after BNP Paribas suspended two subprime mortgage funds), and 03/16/2008 (when JPMorgan Chase acquired Bear Stearns for less than 7% of its market value two days before). The model is estimated separately for each corresponding period. I compute 95% confidence intervals (light blue dotted lines) using the bootstrap.

Table 1 contains the estimates of the upper bounds for the exposure and signaling components in Equation (25). It reports the averages for each bank and each period. For most banks, signaling seems to be the main driver of systematic misreporting, in all five periods. Period (iii) displays the largest deviations. In this period, signaling alone is responsible for differences of approximately -20 bp, on average, and this number is only an upper bound. In contrast, during the same period, with a few notable exceptions, the absolute differences due to exposures are lower than 6 bp. For all other periods, the magnitudes are much smaller. In periods (i) and (ii), the largest estimated deviation due to signaling is -3.5 bp, and the absolute differences attributable to exposures are all below 1 bp. Finally, in periods (iv) and (v) misreporting due to signaling remains high compared with that before Lehman's bankruptcy.

Overall, the model attributes most of the difference between the published Libor and the unmanipulated Libor to signaling rather than exposure to Libor. Figure 6 shows the spread between Libor and the lower bound on the unmanipulated Libor obtained from a model that only allows for exposure as a motive for misreporting (I set $v_{i,p}$ to zero for all banks and periods and then estimate the model). In that model, the estimated Libor manipulation ranges from -8 to 4 bp from 09/03/2007 to 12/31/2010, which is broadly consistent with the results obtained in previous studies that have focused exclusively on this incentive (Snider and Youle, 2012; and Youle, 2014). However, when signaling is added to the model the implied deviations are substantially larger in all five periods, and signaling accounts for most of the resulting spread. To evaluate whether this result depends on Assumption 8, in Section 6.4, I report an alternative estimation in which I allow for persistent differences in idiosyncratic costs across banks. Reassuringly, the result appear robust to dropping this assumption.

Bank	(i): 09/03/2007 - 03/14/2008 Signaling	Exposure	(ii): 03/15/2008 - 09/14/2008 Signaling	Exposure	(iii): 09/15/2008 - 12/31/2008 Signaling	Exposure	(iv): 02/01/2009 - 12/31/2009 Signaling	Exposure	(v): 01/01/2010 - 12/31/2010 Signaling	Exposure
Barclays	0	0.8*	0	-0.1*	0	5*	-6*	-1.2*	-5.5*	-2.5*
BOA	-2.1*	0.1	-2.9*	-0.4*	-27.3*	0.1	-7.2*	-0.6	-8.5*	0.3*
BTMU	-0.7	-0.4*	-0.8*	0.1	-16*	5.3	-0.8	-0.5	-3.1*	-0.4*
Citi	-3.3*	-0.5*	-2.6*	0	-22.8*	-8.1	-9.7*	0.5	-9.3*	0.2
Credit Suisse	-1.1*	-0.2	-0.5*	-0.6*	-13.1*	1.9	-5.1*	-1.2	-5.3*	-1.2*
Deutsche	-2.8*	0.2*	-1.2*	-0.4*	-16.9*	-2.3	-7.2*	-3.7*	-9.7*	-0.4
HBOS	-0.7*	-0.5*	-0.6*	-0.7*	-6*	-5.8*	-	-	-	-
HSBC	-3.5*	-0.4*	-2.2*	0	-34.6*	5.9*	-9.4*	-1.2	-8.2*	-2.3*
JPM	-2.4*	0.2*	-3.5*	0.1*	-47.4*	-0.7	-11.9*	-0.3	-9.8*	-0.4
Lloyds	-2.3*	0.6*	-2.8*	0.2*	-25.2*	-5.5	-7.3*	0	N/A	N/A
Norinchuckin	-0.4	0.1	-0.7*	0.1	-1.1	-14.7*	0	-0.7	-3*	0.3*
Rabobank	-3.5*	-0.6*	-2.4*	-0.3	-34.8*	-3.7*	-10.5*	0.9	-9.6*	0.6*
RBC	-1.9*	-0.1	-1.5*	-0.5	-23.4*	0.4	-0.2	-3.9*	-2*	-4.9*
RBS	-2.4*	-0.2	-2*	-0.1	-10.6*	-2.1	-0.5	-0.3	-4.8*	-1.4*
Societe Generale	-	-	-	-	-	-	-8.5*	-0.1	-26.3*	20.2*
UBS	-3.4*	0.1*	-0.9*	-0.8*	-26.8*	4.5	-4.1*	-1.8*	-9.7*	0.2
WestLB	-1.3*	-0.1	-0.7	-0.2	-22.5*	1	-1.1	0.2	0	0

Table 1: Signaling vs Exposure Incentives

This table presents estimates of two separate components of the average differences $s_i - r_i$, that depend on whether incentives to misreport are driven by signaling or exposures to Libor. Signaling is defined as $\frac{v_{i,p}}{\tilde{N}} \phi_{i,p}\left(r_{i,t,p}|c_{t,p}\right) - v_{i,p}$, and exposure as $\frac{\alpha_{i,p}}{\tilde{N}} \phi_{i,p}\left(r_{i,t,p}|c_{t,p}\right)$. All the estimates reported are upper bounds for the corresponding parameters of the model, measured in basis points. In January, 2009, Societe Generale replaced HBOS as a new member of the USD Libor panel.
*Significant at the 5% level (inference is performed using the bootstrap, as described in Online Appendix B).

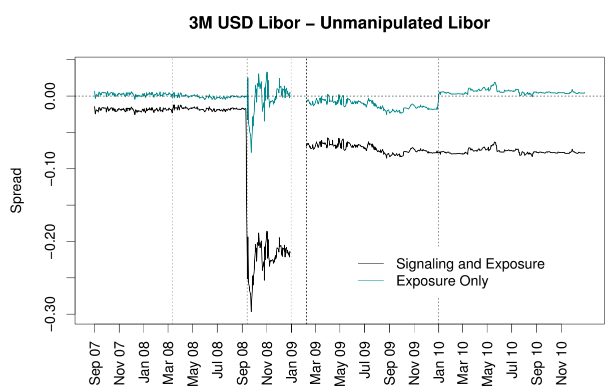

Figure 6: Spread between Libor and the Estimated Common Borrowing Costs with no Signaling

This figure shows the spread between Libor and the lower bound on the unmanipulated Libor obtained from a model that only allows for exposure incentives (I set $v_{i,p}$ to zero for all banks and periods and re-estimate the model). For comparison, I also include the spread estimated with the baseline model. The highlighted dates are: 03/16/2008, 09/15/2008, 12/31/2008, 02/09/2009, and 12/31/2009. Data on Libor individual reports from 01/01/2009 to 02/09/2009 are not available.

In the baseline model, estimated misreporting due to exposure shows substantial variation. This is not very surprising. Most likely, exposures to Libor differ considerably across banks and change over time as outstanding contracts mature and new positions are opened. In contrast, the ranking of banks' signaling incentives tends to persist for longer periods. The correlation coefficient between the vectors of the signaling parameters for two consecutive periods range from 0.61 to 0.77. However, the magnitudes of these parameters substantially differ across banks and over time. Recall, however, that the normalization of the payoff function in (3) implies that parameters $v_{i,p}$ and $\alpha_{i,p}$ measure the gains from misreporting as a fraction of the expected costs from misreporting. Hence, a possible contributor to the overall heterogeneity in these estimates are changes to the likelihood of being caught misreporting driven, for instance, by swings in market volatility. Another potential explanation for the variation in signaling motives is a difference in market perceptions of credit risk. I explore this hypothesis further in Section 6.3.

6.3 Signaling and CDS Spreads

As discussed in Section 2, CDS spreads for bonds issued by banks in the USD Libor panel increased dramatically during the financial crisis and remained high during the subsequent sovereign debt crisis in Europe. The higher perception of credit risk reflected in CDS spreads might have induced banks to use their Libor submissions to try to alleviate concerns about their creditworthiness. Consistent with this hypothesis, I find virtually no indication of manipulation before the BNP Paribas announcement, when CDS spreads were much lower than during the crisis (see Figure 1). To further explore this association, I regress the the estimated signaling parameter $v_{i,p}$ on daily CDS spreads, and control for the bank-specific TED spread. The TED spread is defined as the difference between the USD three-month Libor and the three-month Treasury Bill yield and is widely used by market participants as a measure of credit risk.

Overall, I find a positive association between credit risk and signaling during the periods (i) - (v), as shown in Table 2.[25] An increase of 1 bp in the the CDS spread is associated with a statistically significant increase of 0.8 - 3.1 bp in the signaling parameter. A large fraction of this association seems to be explained by co-movements in time, for any given bank, rather than by cross-sectional variation. When controlling for bank fixed effects, the regression coefficient increases to 3.08 bp and is highly significant. However, when I also include period fixed effects, the coefficient on the signaling parameter drops to 0.75 bp and is no longer significant at the 5% level (the respective p-value is 0.06, with standard errors clustered by bank to control for autocorrelation). Overall, differences in CDS spreads seem to account for only a modest fraction of the cross-sectional heterogeneity in signaling incentives.

As explained in Section 3.3, $v_{i,p}$ is measured as a fraction of the expected cost of misreporting $\gamma_{i,p}$.

[25] The positive association between signaling and CDS spreads reported in Table 2 might explain why Abrantes-Metz, Kraten, Metz, and Seow (2012) do not find a positive correlation between the cross-sectional ordinal rankings of Libor submissions and CDS spreads. Absent any misreporting, they should be positively correlated because both reflect credit risk, but provided signaling incentives, banks with high CDS spreads would understate their borrowing costs.

	(1)	(2)	(3)
CDS Spread$_{i,t}$	2.51***	3.08***	0.75*
	(0.63)	(0.88)	(0.42)
TED Spread$_{i,t}$	1.50	2.40***	-1.16***
	(1.16)	(0.88)	(0.31)
Period FE	No	No	Yes
Bank FE	No	Yes	Yes
Adjusted R^2	0.07	0.65	0.84

Table 2: Regression of Signaling Parameter on CDS Spread

The dependent variable is the estimated signaling parameter $v_{i,p}$. CDS Spread$_{i,t}$ is the daily spread for five-year credit default swaps (CDS) on bonds issued by bank i. TED Spread$_{i,t}$ is the difference between the 3M USD Libor quote of bank i and the yield on 3M US Treasury bills on day t. The sample encompasses 09/03/2007 to 12/31/2010 and includes only banks in the USD Libor panel with available data on CDS spreads. Standard errors (in parentheses) are clustered at the bank level to control for autocorrelation. *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Hence, some of the heterogeneity in the estimated signaling parameters may be due to differences in the expected costs of misreporting that are not related to credit risk. Moreover, as mentioned in Section 4.3, Assumption 8 might also result in cross-sectional heterogeneity in $v_{i,\cdot}$. If the model does not allow the means of the idiosyncratic costs to differ across banks, then it attributes the differences in average reported rates that are not explained by heterogeneous exposures to differences in signaling. The model tends to assign higher signaling parameters to banks with lower reported rates. If lower CDS spreads are also associated with persistently lower idiosyncratic borrowing costs, then the results in Table 2 underestimate the association between CDS spreads and signaling. In particular, the coefficient may be substantially attenuated when period fixed effects are included. Finally, some of the cross-sectional estimated heterogeneity is likely attributable to the variance of estimators of the the bank-specific parameters, and this may be of particular concern in period (iii) because of the small sample size.

6.4 Robustness Check: Heterogeneous Private Costs Means

In this section I report the results of an alternative estimation in which I drop the assumption that the mean of the private costs distribution is the same for all banks. Instead, I estimate the mean for each bank but now I restrict the signaling parameter v to be the same for all banks. This version of the model requires an additional measure of the common cost component c_t to separately identify signaling and exposure to Libor as the drivers for misreporting. Following Kuo et al. (2012), I use the NYFR and the EDR as rough approximations of the common cost (see Section 2). In both cases, signaling seems to be the main driver of misreporting, which confirms the result in Table 1. Thus, the former does not depend on the assumption about the means of the private costs.

Table 3 presents the results of the alternative estimation. The table shows the average absolute misreporting within a period and across all banks ($\frac{1}{N}\frac{1}{T_p}\sum_{i,t}|r_{i,t,p} - s_{i,t,p}|$) that the model attributes to signaling and exposure to Libor according to the decomposition in (25). In period (iii) signaling alone is responsible for absolute differences of, approximately, 14 to 90 bp between bank's borrowing

	(i): 09/03/2007 - 03/14/2008		(ii): 03/15/2008 - 09/14/2008		(iii): 09/15/2008 - 12/31/2008		(iv): 02/01/2009 - 12/31/2009		(v): 01/01/2010 - 12/31/2010	
	Sig.	Expo.	Sig.	Expo.	Sig.	Expo.	Sig.	Expo.	Sig.	Expo.
EDR	5.39	0.4	17.47	1.27	92.72	6.66	29.35	2.44	11.06	2.61
NYFR	-	-	1.01	0.29	14.62	4.56	0.09	1.06	2.13	2.58

Table 3: Average differences between Libor Quotes and Borrowing Costs

This table presents the average absolute misreporting ($\frac{1}{N}\frac{1}{T_p}\sum_i\sum_t |r_{i,t,p} - s_{i,t,p}|$) due to signaling and exposure to Libor. For this robustness check, the means of the private cross distributions are allowed to differ across banks and the signaling parameter $v_{i,p}$ is the same for all banks in each period p. The estimates are based on the assumption that either the Eurodollar deposits rate (EDR) or the NewYork Funding Rate (NYFR) are approximate measures of the average borrowing costs, c_t. The NYFR was introduced in June 2, 2008. I report individual estimates for each bank in Online Appendix B.

costs and their quotes. In contrast, the average absolute difference due to exposure is lower than 7 bp. There are substantial differences depending on whether the EDR or NYFR is used to approximate c_t, given the large spread between these two rates (see Figure 1). In all other periods, misreporting due to signaling seems much lower than in period (iii), but is still substantially larger than what the model attributes to exposure when based on the EDR. However, if the NYFR is used instead to approximate c_t, the corresponding signaling estimate within period (iv) is virtually zero (although it is 29 bp when based on the EDR). Only for this period, and just for the most conservative estimates (NYFR), the deviations from exposure seem larger in magnitude than those motivated by signaling. The differences across periods also resemble the results presented in Section 6.2. I report individual estimates for each bank in Online Appendix B.

6.5 Variation in Borrowing Costs

The estimation results allow us to decompose the variation in borrowing rates into common and idiosyncratic factors. Only for this decomposition, I further assume that the common cost follows a random walk, $c_{t,p} = c_{t-1,p} + \eta_{t,p}$ (an augmented Dickey-Fuller test does not reject a unit root, as reported in Online Appendix B). In Table 4, I compare the estimated standard deviation of $\eta_{t,p}$ to the mean, lowest and largest estimated standard deviations of the private costs $\pi_{i,t,p}$ across banks, for each period p. I report individual estimates for each bank in Online Appendix B.

In periods (i) and (ii), after the BNP Paribas announcement but before Lehman's failure, aggregate shocks seem to account for most of the variation in interbank funding rates for almost all banks. Lehman's collapse disrupted the market severely. Consequently, the volatility of both aggregate and idiosyncratic shocks increases dramatically in period (iii), most likely reflecting much higher uncertainty about idiosyncratic and systemic defaults. There is also much more heterogeneity across banks during this period as reflected by the gap between the lowest (2pb) and the largest (30pb) idiosyncratic factors. Periods (iv) and (v) display the lowest standard deviation of aggregate shocks (1.9 and 0.5 bp, respectively). However, on average, idiosyncratic costs in period (iv) remain more volatile than in periods (i) and (ii), and have higher standard deviations that the common shocks. This could be the result of the declining market activity for unsecured interbank funds during the

	(i): 09/03/2007 - 03/14/2008	(ii): 03/15/2008 - 09/14/2008	(iii): 09/15/2008 - 12/31/2008	(iv): 02/01/2009 - 12/31/2009	(v): 01/01/2010 - 12/31/2010
$sd\,(c_{t,p} - c_{t-1,p})$	5.5	2	11.4	1.9	0.5
mean $\sigma_{\pi_{itp}}$	1.7	1	11	3.5	2
min $\sigma_{\pi_{itp}}$	0.5	0.5	2.3	1.4	0.4
max $\sigma_{\pi_{itp}}$	3.6	2.3	29.9	6.7	9.9

Table 4: Standard Deviation of Common and Idiosyncratic Costs
This table presents estimates of the standard deviation of the common cost component $c_{t,p}$, and the mean, minimum, and maximum standard deviation of the private costs $\pi_{i,t,p}$ across banks. $c_{t,p}$ is assumed to follow a random walk. All estimates are measured in basis points.

crisis. As banks have fewer transactions to base their submissions on, their estimates of their own borrowing costs become more noisy.

7 The end of Libor

The Libor manipulation scandal triggered regulatory reforms around the world to replace interest rate benchmarks that are prone to manipulation. Libor regulators in the United Kingdom announced in March 2021 that the publication of the USD Libor would cease on June 30, 2023. In the United States, financial regulators and the federal government have recently approved laws and regulations to replace the USD Libor with the Secured Overnight Financing Rate (SOFR) which reflects the cost of secured overnight funding in the U.S. Treasury repo market.[26] In contrast to Libor, it is entirely based on data on a large volume of daily transactions. In Europe, the most widely used benchmarks before reform were EONIA and Euribor. Recently, the ECB developed the euro short-term rate (ESTR) as a replacement for EONIA. The ESTR reflects the costs of unsecured overnight borrowing for banks in the euro area. It is exclusively based on deposit transactions that a sample of banks report following ECB regulations. Euribor is not scheduled to be replaced but has recently been reformed to comply with current regulations. It is now based on the daily contributions of a panel of 18 banks. Individual contributions are the weighted average of eligible transactions in the unsecured euro money market and the benchmark is a trimmed mean of these contributions.

In none of these three cases are the contributions or transactions of each bank published, which likely strongly attenuates the signaling motive for manipulating the rate. However, for benchmarks that are based on unsecured transactions and, hence, reflect credit risk, it is still possible that regulators use the data to assess individual credit risk, especially during times of financial distress. Moreover, whether the transactions used to compute such benchmark should be publicly available, including the identities of the counterparties involved, remains an open question. Ultimately, the question is whether transparency is a determinant of the benchmark quality.

Even if the signaling incentive is fully removed, Duffie and Dworczak (2021) have noted that

[26]On March 15, 2022, the U.S. government enacted the Adjustable Interest Rate (LIBOR) Act to establish a process for replacing LIBOR with SOFR in existing contracts that lack provisions for Libor cessation.

transaction-based benchmarks do not eliminate all the incentives or the ability to manipulate them. For instance, banks with long exposure to a benchmark may borrow at rates above their marginal valuation if the benefits from distorting the benchmark more than offset the cost of overpaying. This becomes a more pressing concern when a handful of banks control a large share of the volume of transactions underlying the benchmark, which is the case for the ESTR and Euribor. Nevertheless, manipulating transactions is likely to be much more costly than misreporting borrowing costs; hence, it is safe to assume that the new benchmarks are less vulnerable to manipulation than Libor.[27]

The model in Section 3 can be used to approximate the anticipated reduction in manipulation from replacing Libor with a transaction-based benchmark. Two main goals of the new benchmarks are to remove the signaling motive and to substantially increase the costs of misreporting or manipulation. Consequently, I conduct a counterfactual exercise in which I estimate the reduction in Libor manipulation that would have resulted from a substantial increase in the cost of misreporting or from fully eliminating the signaling incentive.[28] I use the estimated borrowing costs for each bank and each period, and simulate the equilibrium of the Libor reporting game under alternative sets of parameters reflecting counterfactual scenarios. To measure the extent of Libor manipulation in each scenario, I calculate the spread between the reported Libor implied by the simulated equilibrium and my estimated unmanipulated Libor. Figure 7 shows the extent of (counterfactual) manipulation under three scenarios. 1) The baseline scenario which is simulated using the values of the parameters estimated from the data (reported in Section 6). 2) A counterfactual scenario with misreporting costs ten times higher than that in the baseline. 3) A counterfactual scenario with no signaling incentives but the same exposure incentive as in the baseline. According to these simulations, both counterfactual policies would have implied a substantial reduction in Libor manipulation relative to the baseline. In period (iii), the average absolute spread in the baseline scenario is approximately 21 bp, but it drops to 2 bp in the counterfactual scenario with high misreporting costs, and 1 bp when the signaling incentive is completely eliminated. In contrast, in period (v), the equilibrium with no signaling implies a positive average spread of 1 bp, while high misreporting costs result in virtually no manipulation of the benchmark. In summary, transaction-based benchmarks are likely to be much less prone to manipulation than Libor, but non-negligible manipulation may still ocurr in periods of high uncertainty. Regulators should be particularly mindful of the risk of restoring the signaling motive if they plan to use the data underlying the benchmark to assess individual risk.

[27] Duffie et al. (2013),Duffie and Stein (2015), Hou and Skeie (2016), and Coulter et al. (2017) have also discussed the reform of interest rate benchmarks.

[28] With a transaction-based benchmark, the payoff function of bank i can still be captured by Equation 1. In such case, $r_{i,t}$ is a summary statistic (a scalar function) of all bank i's transactions that may be used to calculate the benchmark rate $\tilde{r}_t$, and $s_{i,t}$ is the same summary statistic of bank i's true borrowing costs (absent any manipulation attempts). $\gamma\left(s_{i,t} - r_{i,t}\right)^2$ is the cost of manipulating transactions, and $\alpha_i \tilde{r}_t$ is the gross benefit from exposure. The equilibrium strategies, however, depend on the method used to calculate the benchmark. For a benchmark rate that is linear on a subset of individual contributions $\{r_{i,t} : i = 1, ..., N\}$, like the reformed Euribor, the first order condition of bank i can be expressed in terms of the probability ϕ_i, as in 6.

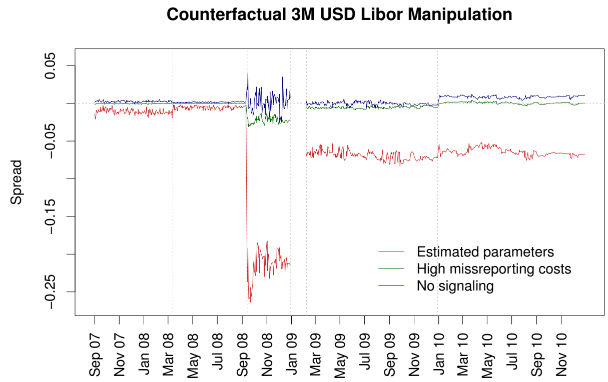

Figure 7: Spread between Libor (simulated under counterfactual scenarios) and the estimated unmanipulated Libor

This figure shows the spread between a simulated Libor (under three different counterfactual scenarios) and the estimated lower bound on the unmanipulated Libor. The highlighted dates are, 03/16/2008, when JPMorgan Chase acquired Bear Stearns for less than 7% of its market value two days before; 09/15/2008, when Lehman Brothers filed for bankruptcy protection; 12/31/2008, when the British Banking Authority changed the composition of the Libor panel; 02/09/2009; and 12/31/2009. Data on Libor individual reports from 01/01/2009 to 02/09/2009 are not available. On 09/24/2008, Libor is below the estimated common borrowing costs of banks in the USD Libor panel by more than 30 bp. 95% confidence intervals (light blue dotted lines) are computed using the bootstrap.

8 Conclusion

I extend structural econometric methods from the empirical auctions literature to estimate a strategic reporting model that identifies signaling and net exposures to Libor as distinct motives for misreporting Libor, using data on banks' Libor submissions only. The model also identifies a lower bound on the unmanipulated Libor, defined here as the value that Libor would have had in the absence of misreporting.

The estimated bound is above Libor during the financial crisis of 2007 - 2008 and the subsequent sovereign debt crisis in Europe, suggesting that Libor understated interbank borrowing costs, which is consistent with the results of previous studies. The estimation results also indicate that signaling was the main driver of misreporting. Regulators concerned with replacing and reforming interest rate benchmarks should pay particular attention to the signaling motive, specifically if data used to compute these benchmarks is also used to assess individual bank risk.

Data Availability Statement

Replication package available at: https://doi.org/10.5281/zenodo.7686525

Data by source

Bloomberg L.P.: access via university terminal subscription

Libor submissions provided by banks in the 3-month USD Libor panel from 09/03/2005 to 12/31/2010. The Bloomberg symbols are: US2003M, US6703M, US0303M, US1403M, US1603M, US2803M, US3403M, US5603M, US7503M, US0103M, US7103M, US1903M, US4803M, US1803M, US3303M, USBN03M, US7603M, USSM03M, US9903M, US2303M, US2203M.

3-month USD Libor from 09/03/2005 to 12/31/2010. The Bloomberg symbol is: US0003M.

New York Funding Rate from 06/02/2008 to 12/31/2010. The Bloomberg symbol is: NYFR3M.

Board of Governors of the Federal Reserve System: Freely available online at FRED

U.S. Short-Term Interest Rates: Daily 3-Month Eurodollar Deposit Rate (DED3) 2005 – 2010

3-Month Treasury Bill Secondary Market Rate, Discount Basis (DTB3) 2005 – 2010

TED Spread (TEDRATE) 2005 – 2010

IHS Markit: access via Wharton Research Data Services (WRDS)

Credit Default Swaps Spreads for 5-year senior unsecured debt issued by banks in the USD Libor panel 2005–2010. Markit RedCode: 06DABK, 0G655D, JJ4650, 189BFD, HK9FHL, 2H6677, 4G425R, 4I75AU, 4C933G, GLA86Z, 6BB62B, NP4897, NUD88R, 8B69AP, HPHB2J, DMFCCI, Tier: SNRFOR, Ccy: EUR, DocClause: CR.

References

Abrantes-Metz, R., S. Villas-Boas, and G. Judge (2011): "Tracking the Libor rate," *Applied Economics Letters*, 18, 893–899.

Abrantes-Metz, R. M., M. Kraten, A. D. Metz, and G. S. Seow (2012): "Libor Manipulation?" *Journal of Banking & Finance*, 36, 136–150.

Athey, S. (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69, 861–89.

Blanco, R., S. Brennan, and I. W. Marsh (2005): "An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps," *The Journal of Finance*, 60, 2255–2281.

Campo, S. (2012): "Risk aversion and asymmetry in procurement auctions: Identification, estimation and application to construction procurements," *Journal of Econometrics*, 168, 96–107.

Cassola, N., A. Hortaçsu, and J. Kastl (2013): "The 2007 Subprime Market Crisis Through the Lens of European Central Bank Auctions for Short-Term Funds," *Econometrica*, 81, 1309–1345.

Chen, J. (2013): "LIBOR's Poker: Interbank Borrowing Costs and Strategic Reporting," http://ssrn.com/abstract=2136871.

Coulter, B., J. Shapiro, and P. Zimmerman (2017): "A Mechanism for LIBOR*," *Review of Finance*, 22, 491–520.

Duffie, D. and P. Dworczak (2021): "Robust benchmark design," *Journal of Financial Economics*, 142, 775–802.

Duffie, D., D. Skeie, and J. Vickery (2013): "A Sampling-Window Approach to Transactions-Based Libor Fixing," FRB of New York Staff Report.

Duffie, D. and J. C. Stein (2015): "Reforming LIBOR and Other Financial Market Benchmarks," *Journal of Economic Perspectives*, 29, 191–212.

Gandhi, P., B. Golez, J. C. Jackwerth, and A. Plazzi (2019): "Financial Market Misconduct and Public Enforcement: The Case of Libor Manipulation," *Management Science*, 65, 5268–5289.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68, 525–574.

Haile, P. A., H. Hong, and M. Shum (2003): "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," Working Paper 10105, National Bureau of Economic Research.

Horowitz, J. L. (2009): *Semiparametric and Nonparametric Methods in Econometrics*, Springer-Verlag New York.

Hortaçsu, A. (2000): "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market," Working paper.

Hortaçsu, A. and D. McAdams (2010): "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market," *Journal of Political Economy*, 118, 833 – 865.

Hou, D. and D. Skeie (2016): *LIBOR: Origins, Economics, Crisis, Scandal and Reform*, London: Palgrave Macmillan UK, 1–13.

Kotlarski, I. (1966): "On Some Characterizations of Probability Distributions in Hilbert Spaces," *Annal di Matematica Pura et Applicate*, 74, 129 – 134.

Krasnokutskaya, E. (2011): "Identification and Estimation of Auction Models with Unobserved Heterogeneity," *Review of Economic Studies*, 78, 293–327.

Kuo, D., D. Skeie, and J. Vickery (2012): "A comparison of Libor to other measures of bank borrowing costs," FRB of New York Staff Report.

Li, T., I. Perrigne, and Q. Vuong (2000): "Conditionally independent private information in {OCS} wildcat auctions," *Journal of Econometrics*, 98, 129 – 161.

——— (2002): "Structural Estimation of the Affliated Private Value Auction Model," *RAND Journal of Economics*, 33, 171–193.

Li, T. and Q. Vuong (1998): "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, 65, 139 – 165.

McAdams, D. (2003): "Isotone Equilibrium in Games of Incomplete Information," *Econometrica*, 71, 1191–1214.

Milgrom, P. and C. Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62, pp. 157–180.

Mollenkamp, C. and M. Whitehouse (2008): "Study casts doubt on key rate; WSJ analysis suggests banks may have reported flawed interest data for Libor." *The Wall Street Journal*.

Newey, W. K. and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 4, 2111 – 2245.

Rose, A. K. (1988): "Is the Real Interest Rate Stable?" *The Journal of Finance*, 43, pp. 1095–1112.

Schrimpf, A. and V. Sushko (2019): "Beyond LIBOR: a primer on the new benchmark rates," *BIS Quarterly Review*.

Shubber, K. and D. Keohane (2018): "SocGen executives ordered Libor rigging, US prosecutors believed," *Financial Times*.

Snider, C. A. and T. Youle (2012): "The Fix Is in: Detecting Portfolio Driven Manipulation of the Libor," http://ssrn.com/abstract=2189015.

Wheatley, M. (2012): *The wheatley review of libor: Final report*, HM Treasury.

Youle, T. (2014): "How Much Did Manipulation Distort the Libor?" Working Paper.

A Proofs

A.1 Non-decreasing Best Response Correspondences

Let the best response correspondence of bank i be defined by

$$\Gamma_i(s_i) = \underset{r_i \in \mathbb{R}}{\operatorname{argmax}} E\left[\alpha_i \tilde{r}(r_i, R_{-i}) + v_i(\tilde{r}(r_i, R_{-i}) - r_i) - \gamma_i (s_i - r_i)^2\right]$$

We will show now that Γ_i is non-decreasing in the following sense (which is stronger than non-decreasing in the strong set order, as defined in Athey, 2001). For all $s_i, s_i' \in S$. with $s_i < s_i'$, $\max \Gamma_i(s_i) \leq \min \Gamma_i(s_i')$.

Let $s_i \in \mathcal{S}_i(c)$, and $r_i = \max \Gamma_i(s_i)$. By definition, for all $r < r_i$

$$\begin{aligned}
E\left[\alpha_i \tilde{r}(r_i, R_{-i}) + v_i(\tilde{r}(r_i, R_{-i}) - r_i) - \gamma(s_i - r_i)^2\right] &\geq E\left[\alpha_i \tilde{r}(r, R_{-i}) + v(\tilde{r}(r, R_{-i}) - r) - \gamma(s_i - r)^2\right] \\
&\Longleftrightarrow \\
(\alpha_i + v_i)(E[\tilde{r}(r_i, R_{-i})] - E[\tilde{r}(r, R_{-i})]) &\geq v_i(r_i - r) + 2\gamma s_i(r - r_i) + \gamma_i(r_i^2 - r^2) \\
&\Longleftrightarrow \\
2\gamma_i s_i(r_i - r) &\geq v_i(r_i - r) + \gamma_i(r_i^2 - r^2) - (\alpha_i + v_i)(E[\tilde{r}(r_i, R_{-i})] - E[\tilde{r}(r, R_{-i})])
\end{aligned}$$

Moreover, since $\gamma_i > 0$, for any $s_i' > s_i$, $2\gamma_i s_i'(r_i - r) > 2\gamma_i s_i(r_i - r)$, then

$$\begin{aligned}
2\gamma_i s_i'(r_i - r) &> v_i(r_i - r) + \gamma(r_i^2 - r^2) - (\alpha_i + v)(E[\tilde{r}(r_i, R_{-i})] - E[\tilde{r}(r, R_{-i})]) \\
&\Longleftrightarrow \\
E\left[\alpha_i \tilde{r}(r_i, R_{-i}) + v_i(\tilde{r}(r_i, R_{-i}) - r_i) - \gamma_i(s_i' - r_i)^2\right] &> E\left[\alpha_i \tilde{r}(r, R_{-i}) + v_i(\tilde{r}(r, R_{-i}) - r) - \gamma_i(s_i' - r)^2\right]
\end{aligned}$$

Thus, for all $s_i' > s_i$, r_i yields a strictly higher payoff than any $r < r_i$. Therefore, any element of $\min \Gamma_i(s_i')$ is at least as large as r_i, in particular $\max \Gamma_i(s_i) \leq \min \Gamma_i(s_i')$.

A.2 Proof of Equation (5).

We want to show that $\frac{\partial E_i[\tilde{r}(r_i, R_{-i})|c]}{\partial r_i}$ is equal to the probability that r_i is included in the computation of the reference rate $\tilde{r}$. The probability is computed from the perspective of bank i, that is, when the quotes of all other banks are random variables with known probability distributions (conditional on the common cost component c).

Proof. Let $R_{-i}^{(n)}$ be the n-th order statistic of the vector R_{-i} of quotes submitted by all banks other than i. From the perspective of bank i, $R_{-i}^{(n)}$ is a random variable with known distribution $G_{(n)|c}$. I follow the convention $R_{-i}^{(1)} \leq R_{-i}^{(2)} \leq ... \leq R_{-i}^{(N-1)}$. Notice that, regardless of i's quote, r_i, for all $k \in \{\underline{n} + 1, ..., \bar{n} - 2\}$, $R_{-i}^{(k)}$ is included in the computation of the reference rate $\tilde{r}$. Moreover, whether r_i, $R_{-i}^{(\underline{n})}$ or $R_{-i}^{(\bar{n}-1)}$ is the other quote included, depends on their relative positions. For instance, with probability $1 - G_{(\underline{n})|c}(r_i)$, $R_{-i}^{(\underline{n})} > r_i$, and in such case $R_{-i}^{(\underline{n})}$ would be included.

Therefore, the expected value of the reference rate $\tilde{r}$, when i submits quote r_i can be written as:

$$\begin{aligned} E\left[\tilde{r}\left(r_i\right)|c\right] &= \frac{1}{\tilde{N}}\left(\sum_{k=\underline{n}+1}^{\bar{n}-2} E\left[R_{-i}^{(k)}|c\right] + E\left[R_{-i}^{(\underline{n})}|R_{-i}^{(\underline{n})} > r_i, c\right]\left(1-G_{(\underline{n})|c}\left(r_i\right)\right)\right. \\ &\quad + r_i\left(G_{(\underline{n})|c}\left(r_i\right) - G_{(\bar{n}-1)|c}\left(r_i\right)\right) \\ &\quad \left. + E\left[R_{-i}^{(\bar{n}-1)}|R_{-i}^{(\bar{n}-1)} \leq r_i, c\right] G_{(\bar{n}-1)|c}\left(r_i\right)\right) \end{aligned}$$

Notice that

$$E\left[R_{-i}^{(\underline{n})}|R_{-i}^{(\underline{n})} > r_i, c\right]\left(1-G_{(\underline{n})|c}\left(r_i\right)\right) = \int_{r_i}^{\bar{r}} x g_{(\underline{n})|c}\left(x\right) dx$$

where $g_{(\underline{n})|c}$ is the probability density function of $R_{-i}^{(\underline{n})}$. A similar expression can be easily found for $E\left[R_{-i}^{(\bar{n}-1)}|R_{-i}^{(\bar{n}-1)} \leq r_i, c\right] G_{(\bar{n}-1)|c}\left(r_i\right)$ and thus, it follows from Leibniz integral rule that:

$$\frac{\partial E\left[\tilde{r}\left(r_i\right)|c\right]}{\partial r_i} = \frac{1}{\tilde{N}}\left(G_{(\underline{n})|c}\left(r_i\right) - G_{(\bar{n}-1)|c}\left(r_i\right)\right)$$

where $G_{(\underline{n})|c}\left(r_i\right) - G_{(\bar{n}-1)|c}\left(r_i\right)$ is precisely the probability that r_i is included in the computation of $\tilde{r}$. □

A.3 Non-binding upper bound for the payoff maximization problem

Consider the maximization problem faced by bank i: $\max_{r_i \in [\underline{r}, \bar{r}]} U_i$, where the payoff function can be written as $U_i = E_{R_{-i}}\left[\beta_i \tilde{r}\left(r_i, R_{-i}\right)|c\right] - v_i r_i - \frac{1}{2}\left(s_i - r_i\right)^2$. I will show that the upper bound $\bar{r}$ can be chosen in such a way that it is never binding. (Although the same holds for $\underline{r}$, zero is a natural lower bound for the set of admissible reported rates).

Clearly,

$$\begin{aligned} \frac{\partial U_i}{\partial r_i} &= \beta_i \frac{\partial E_{R_{-i}}\left[\tilde{r}\left(r_i, R_{-i}\right)|c\right]}{\partial r_i} - v_i + s_i - r_i \\ &= \frac{\beta_i}{\tilde{N}} \phi_i\left(r_{i,t}|c_t\right) - v_i + s_i - r_i \end{aligned}$$

where $\phi_i\left(r_{i,t}|c_t\right)$ is the probability that r_i is included in the computation of $\tilde{r}$, as shown above. obviously, $\frac{\partial U_i}{\partial r_i} < 0$ iff $\frac{\beta_i}{\tilde{N}}\phi_i\left(r_{i,t}|c_t\right) - v_i + s_i < r_i$. Moreover, $0 \leq \phi_i\left(r_{i,t}|c_t\right) \leq 1$. Then $\frac{\beta_i}{\tilde{N}}\phi_i\left(r_{i,t}|c_t\right) - v_i + s_i < \frac{|\beta_i|}{\tilde{N}} - v_i + \bar{s}_i$, where $\bar{s}_i$ is the upper bound of $\mathcal{S}_i\left(c\right)$. Therefore, $r_i > \frac{|\beta_i|}{\tilde{N}} - v_i + \bar{s}_i$ implies that $\frac{\partial U_i}{\partial r_i} < 0$. It follows that if $\bar{r}$ is larger than $\frac{|\beta_i|}{\tilde{N}} - v_i + \bar{s}_i$ for all $i \in \mathcal{N}$, the constraint is never binding.

A.4 Proof of Proposition 2

Proposition. *Let $c \neq 0$, if the strategy of all bank $j \neq i$ is*

$$\rho_j (c + \pi_j; c) = c + \rho_j (\pi_j; 0)$$

Then i's best response is $\rho_i (c + \pi_i; c) = c + \rho_i (\pi_i; 0)$.

Proof. Consider the game with $c = 0$ and let $q_i = \rho_i (\pi_i; 0)$. Clearly, for all $\breve{q}_i \in \left[\underline{q}_i, \bar{q}_i\right]$, $u_i (q_i, \pi_i) \geq u_i (\breve{q}_i, \pi_i)$. Then

$$(\alpha_i + v) E [\tilde{r} (q_i, q_{-i}) |0] - vq_i - \gamma (\pi_i - q_i)^2 \geq (\alpha_i + v) E [\tilde{r} (\breve{q}_i, q_{-i}) |0] - v\breve{q}_i - \gamma (\pi_i - \breve{q}_i)^2$$

where q_{-i} is the vector of actions of all other players $j \neq i$ and $q_j = \rho_j (\pi_j; 0)$. For any $c \neq 0$, let $r_j = \rho_j (c + \pi_j; c)$ then, by assumption $r_j = c + q_j$ for all $j \neq i$.

Intuitively, compared to the game where $c = 0$, when $c \neq 0$, from the point of view of i, the distributions of all other banks' quotes are the same, except for a change in a location parameter c. Thus,

$$E [\tilde{r} (\breve{q}_i + c, R_{-i}) |c] = E [\tilde{r} (\breve{q}_i, q_{-i}) |0] + c$$

for all $\breve{q}_i \in \left[\underline{q}_i, \bar{q}_i\right]$, and it follows that

$$\begin{aligned} (\alpha_i + v) E [\tilde{r} (q_i + c, R_{-i}) |c] - v (q_i + c) - \gamma (c + \pi_i - (q_i + c))^2 &= \\ (\alpha_i + v) (E [\tilde{r} (q_i, q_{-i}) |0] + c) - v (q_i + c) - \gamma (\pi_i - q_i)^2 &\geq \\ (\alpha_i + v) (E [\tilde{r} (\breve{q}_i, q_{-i}) |0] + c) - v (\breve{q}_i + c) - \gamma (\pi_i - \breve{q}_i)^2 & \end{aligned}$$

Since every $\breve{r}_i \in \left[\underline{q}_i + c, \bar{q}_i + c\right]$ can be written as $\breve{r}_i = \breve{q}_i + c$ for some $\breve{q}_i \in \left[\underline{q}_i, \bar{q}_i\right]$, it follows that for such $\breve{r}_i$, $u_i (q_i + c, c + \pi_i) \geq u_i (\breve{r}_i, c + \pi_i)$ and, thus, $r_i = q_i + c$ is i's best response when $c \neq 0$, all other banks strategies are $\rho_j (c + \pi_j; c) = c + \rho_j (\pi_j; 0)$ and its cost is $c + \pi_i$. That is, in the game with c, bank i's best response strategy is

$$\rho_i (c + \pi_i; c) = c + \rho_i (\pi_i; 0)$$

□

A.5 Proof of Proposition 3

Proposition. *If the characteristic functions of the normalized quotes $q_{i_1,t}$, $q_{i_2,t}$ and $q_{i_3,t}$ are non-vanishing everywhere, their distributions are identified, up to an additive constant, from the joint distribution of $(r_{i_1,t} - r_{i_2,t}, r_{i_3,t} - r_{i_2,t})$.*

Proof. Notice that (10) implies

$$\begin{aligned} r_{i_1,t} - r_{i_2,t} &= -q_{i_2,t} + q_{i_1,t} \\ r_{i_3,t} - r_{i_2,t} &= -q_{i_2,t} + q_{i_3,t} \end{aligned}$$

where $q_{i_1,t}$, $q_{i_2,t}$ and $q_{i_3,t}$ are mutually independent and, thus, these two equations satisfy the assumptions of Kotlarski's Lemma. Let Ψ denote the joint characteristic function of $(r_{i_1,t} - r_{i_2,t}, r_{i_3,t} - r_{i_2,t})$ and Ψ_1, its partial derivative with respect to its first argument. Also, let $\Phi_{q_{i_1}}$, $\Phi_{-q_{i_2}}$ and $\Phi_{q_{i_3}}$ denote the characteristic functions of $q_{i_1,t}$, $-q_{i_2,t}$ and $q_{i_3,t}$, respectively. Then, as shown in Li and Vuong (1998),

$$\begin{aligned} \Phi_{-q_{i_2}}(s) &= \exp\left(\int_0^s \frac{\Psi_1(0,u)}{\Psi(0,u)} du - isE[q_{i_1}]\right) \\ \Phi_{q_{i_1}}(s) &= \frac{\Psi(s,0)}{\Phi_{-q_{i_2}}(s)} \\ \Phi_{q_{i_3}}(s) &= \frac{\Psi(0,s)}{\Phi_{-q_{i_2}}(s)} \end{aligned} \tag{26}$$

Since the characteristic function of a random variable uniquely determines its distribution, it follows that the distributions of $q_{i_1} - E[q_{i_1}]$, $q_{i_2} - E[q_{i_1}]$ and $q_{i_3} - E[q_{i_1}]$ are all identified. □

A.6 Proof of Proposition 4

Proposition. *Under Assumptions 2-6, the preference parameter β_i is identified, for all $i \in \mathcal{N}$, from the distributions of the normalized quotes. Moreover, if $\tilde{q}_i = q_i - E[q_{i_1}]$,*

$$\beta_i = \frac{Med(\tilde{q}_i) - E[\tilde{q}_i]}{\phi_i(Med(\tilde{q}_i)| - E[q_{i_1}]) - E[\phi_i(\tilde{q}_i| - E[q_{i_1}])]}$$

Proof. Notice that Equation (15) can be interpreted as the inverse equilibrium strategy of bank i, in the game with $c = -E[q_{i_1}]$ (translated by a constant). Hence, it implicitly defines the equilibrium strategy in the interior of $\mathcal{S}^{\mathcal{I}}(c)$, the support of π_i. By assumption, $E[\pi_i]$ is an interior point in $\mathcal{S}^{\mathcal{I}}(c)$.

Let q_i^* be the optimal quote when $\pi_i = E[\pi_i]$. Since q_i^* is an interior solution, then:

$$q_i^* - (E[q_i - E[q_{i_1}]]) - \beta_i(\phi_i(q_i^*| - E[q_{i_1}]) - E[\phi_i(q_i^*| - E[q_{i_1}])]) = E[\pi_i] \tag{27}$$

Let $\varrho(q) = q - (E[q_i - E[q_{i_1}]]) - \beta_i(\phi_i(q| - E[q_{i_1}]) - E[\phi_i(q| - E[q_{i_1}])]) - E[\pi_i]$. Since the equilibrium strategy is strictly increasing in $\mathcal{S}^{\mathcal{I}}(c)$

$$\varrho(\tilde{q}_i) < 0 \text{ for all optimal } \tilde{q}_i < q_i^* \quad \text{and} \quad \varrho(\tilde{q}_i) > 0 \text{ for all optimal } \tilde{q}_i > q_i^*$$

It follows that $P_{\tilde{q}_i}\{\tilde{q}_i \leq q_i^*\} = P_{\tilde{q}_i}\{\varrho(\tilde{q}_i) \leq 0\} = P\{\pi_i \leq E[\pi_i]\}$. Since, by assumption, $P\{\pi_i \leq E[\pi_i]\} =$

$\frac{1}{2}$, then $P\{q \leq q_i^*\} = \frac{1}{2}$ and hence q_i^* is the median of the distribution of $\tilde{q}_i$. Let $\text{Med}(X)$ denote the median of X. We just showed that $\text{Med}(\pi_i) = E[\pi_i]$ implies $\text{Med}(q_i) = q_i^*$. Besides, q_i^* satisfies Equation (27), which uniquely identifies β_i as

$$\beta_i = \frac{\text{Med}(\tilde{q}_i) - E[\tilde{q}_i]}{\phi_i(\text{Med}(\tilde{q}_i)| - E[q_{i_1}]) - E[\phi_i(\tilde{q}_i| - E[q_{i_1}])]}$$

□

A.7 Variance of the estimated normalized quotes

Equation (19) can be used to obtain an estimator of $Var[q_{i,t}]$, for all $i \in \mathcal{N}$, as follows.

$$Var[r_{i,t} - \bar{r}_t] = \left(\frac{N-1}{N}\right)^2 Var[q_{i,t}] + \frac{1}{N^2}\sum_{j \neq i} Var[q_{j,t}] \tag{28}$$

Since $Var[r_{i,t} - \bar{r}_t]$ can be estimated directly from the data, (28) provides a system of linear equations that can be solved for $Var[q_{i,t}]$, for all $i \in \mathcal{N}$. Let $\hat{\sigma}^2_{q_i}$ denote the corresponding consistent estimator of $Var[q_{i,t}]$. Also, let $\bar{r}_t^w = \sum_{i \in \mathcal{N}} \frac{r_{i,t}}{\hat{\sigma}^2_{q_i}} / \sum_{i \in \mathcal{N}} \frac{1}{\hat{\sigma}^2_{q_i}}$ and $\bar{q}_t^w$ be similarly defined (using the same weights). Moreover, let $\xi_t^w = \bar{q}_t^w - E[\bar{q}^w]$. It follows that $r_{i,t} - \bar{r}_t^w = q_{i,t} - E[\bar{q}^w] - \xi_t^w$. Then $\hat{Var}[\xi_t^w] = \left(\sum_{i \in \mathcal{N}} \frac{1}{\hat{\sigma}^2_{q_i}}\right)^{-1}$ provides an estimate of $Var[\xi_t^w]$.