Bank funding and risk taking*

Alessandro Ferrari, Carmen García Galindo, Matic Petricek,
Andreas Winkler†

First Version: May 2017

December 5, 2018

Abstract

In this paper we use a novel approach to address issues of endogeneity in estimating a causal effect of leverage on risk taking by banks. Using data on local bank office deposits and local unemployment we construct an instrument to use in a regression of leverage on a measure of risk taking constructed from new issuance of loans. The results confirm that banks increase their risk taking after an exogenous increase in leverage.

JEL Codes: G11, G20, G21, G28

1 Introduction

There are two established, opposing theoretical results about the effect of leverage on risk taking by banks. First, due to limited liability, expected returns on equity investment increase with an increased riskiness of the portfolio. As a bank’s equity holders are protected from the left tail of the returns to assets distribution by limited liability, they have an incentive to increase the variance of the distribution by taking on more risk. On the other hand, callable demand deposits constitute a substantial part of bank debt. On average, almost 20 percent of these are above

*We are thankful to Juan Dolado, Ramon Marimon and all the participants to various seminars at EUI for useful comments and suggestions. All the remaining errors are ours.

†European University Institute
the amount insured by the Federal deposit insurance corporation. This provides depositors with a strong incentive to monitor as well as with tools to punish excessive risk taking.

There is inconclusive empirical evidence on which of the two effects prevails. Some authors such as Koudstaal and van Wijnbergen (2012) find that higher leverage leads to less risk taking, while other such as Acosta Smith et al. (2017) find that an increase in capital requirements leads to a decrease in risk taking. Finally, there are some authors such as Jacques and Nigro (1997) that find no effect at all. As we argue below, the results in all these papers are, however, influenced by endogeneity issues. Our aim in this paper is then to estimate the causal effect of leverage on risk taking behaviour of banks. In doing so we add to the literature by proposing a novel approach to addressing endogeneity in this setting.

We identify two sources of endogeneity and propose a method to address them. First, an increase in leverage may incentivise banks to pursue riskier investment, but at the same time the demand and supply of deposits, which drive a bank’s leverage given equity, can also be affected by a bank’s risk taking, giving rise to reverse causality. Such an effect may arise if a bank becomes known for making risky investment choices and is consequently avoided by depositors. Second, shocks, observable or non-observable, common to both assets and liabilities of a bank, if omitted, can cause a bias in the estimate of the effect of leverage on risk taking. This, in particular, is an issue that will arise whenever one measures risk as realised risk in the portfolio, as is common in the previous literature.

We conduct our empirical analysis by using instrumental variable estimation. We first address the issue of reverse causality by making use of bank office level data on deposits for US banks and geographically granular unemployment data. We argue that local unemployment rates are exogenous to the risk taking of a bank as a whole, and construct an instrument based on bank’s exposure to de-
posit supply shocks caused by changes in local unemployment rates, to use in our final regression of leverage on risk. However, as is often the case in the empirical literature, if risk taking is approximated by using a risk measure of existing portfolios, the issue of omitted shocks, which may be common to both assets and liabilities, remains. To address this, we construct a measure of risk taking based solely on newly issued mortgage loans, by considering the universe of mortgage loan applications from 1999 to 2016. We argue that, while the existing portfolio of a bank can be affected by geographical area specific shocks which affect deposits and leverage at the same time, newly issued loans are chosen by banks after the shocks have realized. Hence, the riskiness of newly issued loans is a choice by the bank, unaffected by local area shocks.

Our results confirm that limited liability induces banks to take on more risk after an exogenous increase in leverage, with more leveraged banks being significantly more likely to issue risky mortgage loans. We find that decreasing a bank’s leverage ratio by 1 percentage point (which corresponds to an increase in leverage), will lead to this bank originating a loan of 'average' risk with a 3.8% higher probability. In a second estimation we find that this translates into a 1% increase in the predicted median probability of default of loans issued by this bank.

The remainder of this paper is structured as follows: section 2 provides a review of the relevant literature, sections 3 explains the methodology and the data. Section 4 presents the results, while section 5 discusses their robustness with respect to the measure of risk taking. Finally, section 6 concludes and discusses the policy implications of the findings.

2 Literature Review

There are several theoretical papers on how banks’ funding structure should impact their risk taking. Jensen and Meckling (1976) show that for a firm the decision to take on debt is equivalent to buying a call option from its creditors.
When the debt is due, they can either choose to redeem the bond (buy back the firm, so to speak) or not to. Since the value of such an option is increasing in volatility (see for example Black and Scholes (1973)), firms have an interest to take on higher amount of risk than without debt financing. On the other hand, as Laeven and Levine (2009) point out, requiring banks to hold more capital may not necessarily reduce these incentives if this capital is raised by issuing equity to new shareholders. Simply adding more shareholders with the same incentives may not actually alleviate the issue. Instead, shareholders may decide to make up for the higher cost of capital by taking on even more risky projects in the spirit of Koehn and Santomero (1980). Finally, Diamond and Rajan (2001) show that demand deposits can have a disciplining effect on banks. Our paper contributes to the literature by providing a well-identified answer about the causal effect of leverage on risk taking.

In the empirical literature, there are two ways that previous work has addressed the question of how leverage impacts bank risk taking. The first strain of papers seeks to provide an answer through a direct regression of a measure of leverage on some measure of risk. Altunbas et al. (2007) aim to identify the relationship between leverage and risk by means of a seemingly unrelated regression design that relates changes in capital and risk. They use loan-loss provisions as a proxy for the risk taken on by the bank and find that in their whole sample, banks with a higher equity to asset ratio will take on more risk, while the relationship is negative for the most efficient banks in the sample. Jacques and Nigro (1997) employ an approach that uses a regulatory pressure variable, a measure of how far a bank’s equity holdings are from the regulatory threshold, to identify the effect of leverage on risk, but can not refute the null hypothesis that leverage has no effect on risk taking. Koudstaal and van Wijnbergen (2012) aim to identify the effect of leverage on risk by regressing the standard deviation of returns on assets on lagged leverage while controlling for market volatility. They find that
higher leverage leads to less risk taking, but that this result is entirely driven by low leverage banks. Highly leveraged banks, they find, do not react to changes in leverage. Similarly to the above paper, Shrieves and Dahl (1992) find that banks take on more risk when there is a positive shock to capital by using a simultaneous regression framework. As we will argue in our methodological section below, all these papers have shortcomings in two ways: first, they fail to provide a convincing identification in the sense that leverage cannot be seen as exogenous in any of the above models. Second, as all the papers employ some measure of portfolio risk, they consider a rather noisy measure of risk taking which is also impacted by market conditions, which may in turn be impacting banks leverage decisions.

The second strain of papers in the literature attempts to provide exogenous variation to leverage by evaluating the effect of policies that impact banks ability to leverage out. Laeven and Levine (2009) find that requiring banks with an owner that holds a significant voting share to hold more capital has the effect of reducing risk taking, while the opposite is true for widely held banks. Acosta Smith et al. (2017) find that the introduction of the leverage ratio requirements in the Basel III framework did cause banks to increase their capital holdings and reduce their risk taking. Finally, Ashraf et al. (2016) show that the introduction of risk weighted capital standards led to a reduction of bank portfolio risk in Pakistan. We add to the empirical literature by providing clean identification of the causal effect of leverage on risk taking, both by adding a new instrument for leverage and by using a measure of risk taking (a banks decision to issue certain loans) that is much less likely to be subject to outside influences and previous choices than the portfolio based measure currently used in the literature. While these papers suffer from the endogeneity associated with leverage to a smaller degree, we believe that our identification is superior as we do not need to rely on the assumption that banks did not react to news or rumors of potential policy changes prior to the implementation of the reform. Additionally, all of these papers rely once more
on portfolio based measures of risk, while the present work employs a much more
direct measure of risk taking.

In parallel work to this, Ohlrogge (2017) uses an approach that is similar to
the one taken in this paper and comes to results that confirm our analysis.

Methodologically, our paper is related to a paper by Bartik (1993) that employs
local industry shares to identify the impact of labour supply on wages. This
approach has been recently formalized and discussed by Goldsmith-Pinkham et al.
(2017).

3 Methodology

3.1 Overview

In this section we explain the endogeneity issues that plague the analysis of lever-
age and risk taking, as well as our methodology to tackle them. Before diving
deeper into those issues, some definitions will prove useful in the following discus-
sion. First, we will use the term risk taking (behaviour) as an act of making
new investments (issuing new loans) with different degree of riskiness attached to
them. This is the subject of our analysis. It is important to distinguish it from
the term riskiness of the portfolio, which is defined by the riskiness attached
to loans which have been issued in the past. Variation in riskiness of the portfo-
lio can be caused by both risk taking behaviour and by current and past shocks
absorbed by the portfolio. Although the riskiness of the portfolio is often used to
proxy risk taking behaviour in the literature, the distinction will be important in
understanding the issue of endogeneity and our identification.

Most commonly, the literature defines an increase in leverage as an increase in
debt financing relative to equity financing by a bank. Basel III however defines
a leverage ratio as core equity relative to total assets. For the purpose of easy
application to the most recent regulation, we use the latter as a measure of leverage
in our econometric analysis. Thus, an increase in leverage due to an increase in
debt financing corresponds to a *decrease* in the leverage ratio.

We identify two sources of endogeneity which prevent a causal interpretation of a simple regression of leverage on some commonly used measure of riskiness of the portfolio.

- Simultaneity/reverse causality: For a given level of equity, more deposits may incentivise banks to undertake riskier investments, but the demand and supply of deposits are also affected by the riskiness of the portfolio and a bank’s risk taking behavior.

- Omitted shocks common to the portfolio and the deposits.

To tackle the first source of endogeneity we build an instrument for leverage that is exogenous to a bank’s risk choices. To this end, we use data on deposits at the office level provided by the FDIC. This detailed geographical information on deposits enables us to compute bank exposure to local unemployment variation, which we use as an instrument providing us with variation in leverage exogenous to bank behaviour. This measure, however, can still be endogenous if it relies on the existing portfolio. If some exogenous shocks to economic activity occur, they are likely to not only affect deposits (and leverage) through unemployment but also through the riskiness of the existing portfolio. To tackle this issue, we construct a measure of risk taking based on the new issuance of mortgage loans. We argue that while riskiness of the portfolio may be affected by some shocks which are common to both deposits and assets, the riskiness of new issuances are a choice for banks.

Following the procedure described above, we conduct our empirical analysis by, first, constructing an instrument to assure variation in leverage that is independent of bank risk taking and constructing a measure of risk taking based on newly issued mortgage loans, which is independent of local area shocks. We then regress our instrumented leverage ratio on the riskiness of new issuances in order to estimate
the causal effect of leverage on bank risk taking. The methodology is explained in more detail after a discussion of the data employed.

3.2 Data

There are four main sources of data that we use in the analysis. The bank-office level deposit data from the FDIC, and the local unemployment data from the Bureau of Labour Statistics are used in constructing the instruments. We use the Home Mortgage Disclosure Act data on the universe of mortgage applications from the FFIEC to construct a measure of risk taking. Finally, we add balance sheet data from the FDIC to calculate the leverage ratio in the IV estimation of the effect of leverage on risk taking.

Summary of deposits (FDIC)

Summary of deposits data is annual data on the level of deposits at the bank office level. For every office, for every bank operating in the US and insured by the FDIC, the amount of deposits as of June 31st is reported. Along with the deposit level, the data contains detailed geographic and demographic information for every office as well as an identifier for the owning bank. This identifier is then used to merge this data with balance sheet data also provided by the FDIC.

For the purpose of our analysis, the data for every bank is collapsed at a relevant geographical area level. We will be using the Core Based Statistical Areas. A Core Based Statistical Area (CBSA) consists of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people plus adjacent counties that are socioeconomically tied to the urban center by commuting. Not all counties are a part of a CBSA. Around 10% of all observations come from counties which are not part of any CBSAs adding up to around 5% of all deposits. We aggregate these counties at the state level into CBSA equivalents and brand them as rural state areas. Figure 1 presents some descriptive statistics from the Summary of Deposits data. Two features of the data are noteworthy. First, over our sample period of 1999 to 2015, there is a significant consolidation in the
banking market, with the number of banks decreasing by about 40%, while the number of offices per bank more than double. We take care that this development does not impact our analysis by removing an office sold by bank $i$ to bank $j$ in year $t$ from the sample in years $t$ and $t - 1$, so that the change in deposits between $t - 1$ and $t$ for neither bank $i$ nor bank $j$ is affected by the transfer of ownership of that office.

**Figure 1: Bank offices data**

Local area unemployment statistics (BLS)

Local area unemployment statistics provide monthly data on unemployment at the county level. Since the relevant geographical area in constructing the instruments is the CBSA, we aggregate the statistics to the CBSA level at yearly frequency. Figure 2 presents the descriptive statistics of unemployment rate of the CBSAs in the US within our sample period. This graph shows that there is indeed significant heterogeneity between CBSAs, when it comes to unemployment, providing variation to exploit when constructing our instrument. As we show later, this heterogeneity in unemployment rates will have heterogenous effects on banks, which are differentiated by the geographic composition of their deposit holdings. Finally, the first stage regression of our instrumental variable approach shows that this variation is strongly correlated with leverage at the bank level.
Figure 2: CBSA unemployment statistics

Home mortgage disclosure act data (FFIEC)
The Home Mortgage Disclosure Act (HMDA) mandates that banks above a set threshold of assets issue detailed reports on their mortgage applications, lending and purchases. The reporting is done through the Loan Application Registries (LAR) and includes all mortgage loan applications within a year. Moreover, the registries contain some characteristics of the applicant and potential co-applicant (ethnicity, race, gender, income), as well as characteristics of the loan (amount, type, purpose, rate spread for some, occupancy), the property (type, census tract, etc.), the census tract in which the property is located (income relative to the MSA, minority population, number of housing units, etc.), as well as the action taken by the bank (origination, denial and its reason, sale to an institution like Freddie Mac).

Our construction of the measure of risk taking loosely follows DellAriccia et al. (2012) and relies on the loan to income ratio. The loan to income ratio is computed as the total loan amount in the application over the total gross annual income an institution relied upon in making the credit decision. To add to the methodology

\[ \text{Loan to income ratio} = \frac{\text{Total loan amount}}{\text{Total gross annual income}} \]

\[ \text{Gross annual income is not registered in HMDA due to four possible reasons: (i.) multi-family dwellings, (ii.) income was not registered in the loan purchase documentation, (iii.) loans} \]

10
on a measure of risk taking we also use the data on origination. We define origination as an application which has been accepted and then either originated or refused by the applicant, a purchase of a loan, or a preapproved request. We define a non-origination as an application denied by the bank or a denied prerequisite. We ignore all applications withdrawn by the applicants or applications closed for incompleteness.

The LAR data reports all applications, accepted or rejected. Table 1 provides the statistics on the origination ratio between 2004 and 2012. The share of originated loan applications (see column (1)) decreased from 74% in 2004 to 67% in 2007. In 2009, the origination ratio increased sharply and then gradually increased to 80% in 2012. The sharp increase reflects the crisis, which has decreased the demand for loans and forced the worse potential borrowers out of the market. The remaining pool was of higher quality which increased banks willingness to lend to remaining applicants.

| Year | Origin. ratio | DtI | Empl. hist. | Credit. hist. | Coll. | Dwnpay. | Other |
|------|---------------|-----|-------------|---------------|------|---------|-------|
| 2004 | .744          | .131| .010        | .299          | .113 | .0154   | .432  |
| 2005 | .729          | .122| .0111       | .267          | .122 | .012    | .466  |
| 2006 | .715          | .150| .0129       | .284          | .154 | .017    | .381  |
| 2007 | .677          | .173| .0123       | .272          | .193 | .017    | .332  |
| 2008 | .681          | .205| .0114       | .265          | .250 | .018    | .249  |
| 2009 | .767          | .227| .0129       | .209          | .310 | .020    | .220  |
| 2010 | .776          | .222| .0134       | .202          | .252 | .021    | .288  |
| 2011 | .766          | .214| .013        | .223          | .241 | .021    | .285  |
| 2012 | .794          | .213| .013        | .233          | .218 | .024    | .301  |

Table 1 also reports the shares of prevailing reasons for rejecting a loan. Insufficient collateral (column (5), high debt-to-income (column (2)) and poor credit history (column (4)) explain the bulk of the rejection decisions. The effect of the crisis is evident in the spike of the share of rejections due to insufficient collateral to bank employees, (iv.) loans to non natural persons. These cases are excluded from the estimation as described in the methodology section.
in 2009 when house prices collapsed. We take into account this crisis effect by including time fixed effects in our estimation of the risk measures.

On top of the loan application data, the LAR reports also the information about the loans purchased by banks. Tables 2 provides the statistics about the characteristics of all applications, the accepted applications, the rejected applications and the purchased loans. The table shows that accepted loans have, on average, a lower loan-to-income ratio than rejected loans.

| Year | All mean | All p50 | Accepted mean | Accepted p50 | Rejected mean | Rejected p50 | Purchased mean | Purchased p50 |
|------|----------|---------|---------------|--------------|---------------|--------------|----------------|---------------|
| 2004 | 2.274    | 2.083   | 2.227         | 2.103        | 2.408         | 2.000        | 2.562          | 2.390         |
| 2005 | 2.281    | 2.098   | 2.217         | 2.106        | 2.452         | 2.077        | 2.480          | 2.400         |
| 2006 | 2.188    | 1.974   | 2.109         | 1.951        | 2.385         | 2.038        | 2.322          | 2.233         |
| 2008 | 2.304    | 2.098   | 2.184         | 2.059        | 2.553         | 2.200        | 2.547          | 2.451         |
| 2009 | 2.333    | 2.209   | 2.311         | 2.174        | 2.687         | 2.300        | 2.735          | 2.607         |
| 2010 | 2.430    | 2.243   | 2.420         | 2.211        | 3.060         | 2.372        | 2.539          | 2.427         |
| 2011 | 2.349    | 2.023   | 2.222         | 2.000        | 2.761         | 2.083        | 2.587          | 2.435         |
| 2012 | 2.384    | 2.054   | 2.291         | 2.051        | 2.735         | 2.065        | 2.573          | 2.409         |

Balance sheet data (FDIC)

To construct the leverage measure of banks, we use balance sheet data provided by the FDIC. The data is available at quarterly level and includes income statements as well as several performance ratios.

### 3.3 Exogenous Variation in Deposits: Two Instruments

The aim here is to construct an instrument to assure that the variation in leverage is independent of risk taking. We do so in two ways: i) estimating shocks to local deposits caused by unemployment changes which we then aggregate for each bank by computing the weighted average of these local shocks; ii) or directly computing the weighted average of unemployment changes and using this as an instrument for leverage.

To this end we use data on deposits at the bank-office level, administered by the Federal Deposit Insurance Corporation (FDIC) and the Local Area Unemployment Statistics administered by the Bureau of Labour Statistics. The first dataset
contains yearly information on the level of deposits for all offices of all banks insured by the FDIC, together with the demographic information on the office and the bank which owns it. The second dataset provides monthly unemployment figures at county level. The relevant geographical definition in our analysis is the Core Based Statistical Area.

The rationale behind these instruments for bank level deposit growth rates follows closely Bartik (1993), whose approach has been extensively analysed in Goldsmith-Pinkham et al. (2017). The standard idea behind this approach is that when one is interested in a parameter, say, the elasticity of labour supply, using changes in wages and employment growth rates at local level, one should be concerned with the endogeneity of local employment growth. To solve this issue, the Bartik approach suggests to define an instrument as the local employment growth predicted by interacting local industry employment shares with national industry employment growth rates.

In our setting we follow a similar logic but we apply it to a different level of granularity of the data. Our potentially endogenous object is the deposit growth rate of banks. We therefore build as an instrument the predicted change in deposits for a bank in a given period as the interaction between the bank’s geographical area deposit share and the change in deposits in the geographical area.

We do so in two different ways to avoid any further endogeneity concern or feedback loop between bank and area level deposit changes: i) we predict the change in deposits in a geographical area in a given period based on the change of local unemployment in that period and use this fitted value as our instrument; ii) we use the change in unemployment in the geographical area directly (not using it to predict deposits) as the instrument.

Before discussing the two approaches in further detail, key differences between our strategy and the standard Bartik instruments should be highlighted. As mentioned above, the literature employs this approach to solve endogeneity problems.
In contrast, we use the geographical area shares as a mean of aggregation, not as a solution to endogeneity per se. We adopt as instrumental variables for the leverage the “relevant” changes in local unemployment or deposit supply, where “relevant” is to be read as weighted by geographical area deposit composition. The adoption of this specific aggregation strategy serves two distinct purposes: first and foremost is an appealing way of aggregating geographical area specific changes to the bank level; second it eliminates any further endogeneity concerns.

The two approaches are formalized below.

**Instrumenting deposits**: In building the first instrument we regress the growth rate of total deposits in a geographical area on the change in the local unemployment rate. We brand the fitted values from this model at the geographical area level as shocks to deposit supply at the geographical area level. For each bank in each year we then compute the exposure to this variation in deposit supply as a weighted average of these shocks using the deposits each bank holds in a particular area as weights. This implies the following procedure,

\[
\Delta \text{dep}_{i,t} = \alpha_0 + \gamma_i + \eta_t + \beta \Delta \text{unemp}_{i,t} + \epsilon_{i,t}
\]  

where \(\Delta \text{dep}_{i,t}\) denotes the growth rate of deposits in a geographical area \(i\) in period \(t\), \(\gamma_i\) and \(\eta_t\) denote geographical area and time fixed effects, and \(\Delta \text{unemp}_{i,t}\) denotes a change in unemployment rate in geographical area \(i\) in period \(t\). We call the fitted values from the model above local deposit supply shocks. To compute the exposure of a particular bank in a particular period to these shocks, which will serve as an instrument for leverage in our final estimation, we compute the weighted average of these shocks for every bank, where then we use the deposit this particular bank holds in different areas as weights. For bank \(b\), operating in areas \(i = 1..I\), this implies:

\[
\Delta \hat{\text{dep}}_{b,t} = \frac{\sum_{i=1}^{I} \text{dep}_{b,i,t} \Delta \hat{\text{dep}}_{i,t}}{\sum_{i=1}^{I} \text{dep}_{b,i,t}}
\]  

where \(\text{dep}_{b,i,t}\) denotes the deposits bank \(b\) holds in geographical area \(i\) in period
We use the measure $\Delta dep_{b,t}$ as one of the possible instruments in the final estimation of the effect of leverage on risk taking.

Table 3 presents the estimation results for equation 1. Results, as expected, prove a negative and highly significant effect of changes in unemployment on deposit growth rates at the CBSA level. An increase in unemployment change in a CBSA by one percentage point decreases the deposit growth rate in that area by 0.43 percentage points after controlling for the CBSA and year fixed effects.

**Table 3: First preliminary stage**

|                  | (1)                |
|------------------|--------------------|
| $\Delta \ln$ (deposits) |                    |
| $\Delta$ unemp   | -0.432***          |
| Constant         | 0.0310***          |

|                  |                   |
|------------------|-------------------|
| Time FE          | YES               |
| CBSA FE          | YES               |
| N                | 21450             |
| $R^2$            | 0.014             |

**STANDARD ERRORS IN PARENTHESES**

** p < 0.05, *** p < 0.01, **** p < 0.001

**Direct local unemployment exposure:** In building the alternative instrument we directly estimate the exposure of each bank to changes in local unemployment rates, using, as before, the deposits a bank holds as weights. For bank $b$, operating in areas $i = 1..I$, this exposure, $\Delta exp_{b,t}$, is given by:

$$\Delta exp_{b,t} = \frac{\sum_{i=1}^{I} dep_{b,i,t} \Delta unemp_{i,t}}{\sum_{i=1}^{I} dep_{b,i,t}}$$  \(3\)

where, as before, $dep_{b,i,t}$ denotes the deposits bank $j$ holds in geographical area $i$ in period $t$, and $\Delta unemp_{i,t}$ denotes a change in unemployment rate in geographical area $i$ in period $t$. We use $\Delta exp_{b,t}$ as the second instrument in the final estimation of the effect of leverage on risk taking.

Figure 3 plots the mean of the bank level deposit supply shock, $\hat{dep}_{j,t}$ across time, and the mean bank level exposure to unemployment, $\Delta exp_{b,t}$. The figure reveals a sharp drop in deposit supply coinciding with a spike in unemployment across the US.
The two instruments exploit the same variation of changes in unemployment at the local level. They are not numerically equivalent due to different specifications of the fixed effects. It is also worth noting that the second instrument does not require any estimation since it is only built through aggregation of local areas changes at the bank level. This is relevant because one may be concerned that our first instrument may suffer from generated regressor problems. As we will show later we obtain fairly similar results with the two instruments.

3.4 A Measure of Risk Taking

Remaining endogeneity
The procedure explained above describes constructing a measure of an exogenous change in deposits and a measure of an exposure of banks to changes in local unemployment rates. Both these measures are exogenous to risk taking, but not exogenous to a measure of riskiness of the portfolio. To exemplify the issue, consider a shock to deposits in a certain area as estimated in the previous section. Such a shock is likely to impact the income of depositors. It cannot, however, be excluded to have impacted also the borrowers, private or corporate, in an area, which may or may not have borrowed from the banks operating in that areas. Any measure of risk, which is based on the performance of the existing portfolio might
be subject to this sort of residual endogeneity.

New issuances of loans are not subject to this endogeneity concern since new issuances can only be affected by the existing pool of potential loans. A geographical area shock can affect the existing local pool of borrowers, while it does not affect the entire pool of potential borrowers. New issuance of a loan is a choice for a bank and the riskiness of new issuances proxies risk taking behaviour.

**Creating a measure of risk taking**

To this end we construct a measure of risk taking based on the issuances of new mortgage loans based on the Home Mortgage Disclosure Act (HMDA) dataset, administered by the Federal Financial Institutions Examination Council (FFIEC). It is a yearly dataset on the population of mortgage applications to banks and other mortgage lenders with detailed information on the borrower and loan characteristics. We take the riskiness of new mortgage lending as representative of risk taking on the entire portfolio.

To construct a measure of risk taking behaviour by banks, we estimate the responsiveness of loan issuance of each bank in each year to riskiness of the borrower and the loan. As a measure of riskiness of the loan and the borrower we use the Loan-to-Income (LtI) ratio computed from the HMDA dataset for every loan application. This follows loosely DellAriccia et al. (2012), where LtI is used directly as a measure of risk in their analysis of lending standards. This methodology
implies the following model:

\[ \text{Origin}_{t,b,j} = \gamma_0^t + \gamma_1^t \text{LtI}_{t,b,j} + \epsilon_{t,b,j} \] (4)

where \( \text{Origin}_{t,b,j} \) denotes a binary loan origination variable which takes the value \( \text{Origin}_{t,b,j} = 1 \) if the application in period \( t \) to a bank \( b \) by a borrower \( j \) is accepted and loan is originated, and takes the value \( \text{Origin}_{t,b,j} = 0 \) if the application is rejected and the loan is not originated. \( \gamma_0^t \) captures the effect of the macroeconomic situation in period \( t \) for all banks, such as market conditions and regulation, which may cause banks to have differing appetites for risk over time. Finally, for every bank \( b \) in every period \( t \) we also obtain an estimate of the risk responsiveness \( \gamma_1^t \) based on Loan-to-Income of all applicants \( j \), which serves as a measure of risk taking behavior by banks.

Figure 4 plots the risk measure for the banks included in the analysis over the years. The distribution has a mean of \(-.0039622\).

**Figure 4: Responsiveness of origination to risk**

\[ \text{Weighted average elasticity to risk} \]

\[ \begin{array}{c}
\text{year} \\
2000 & 2005 & 2010 & 2015 \\
\text{Weighted average elasticity to risk} & \end{array} \]

---

\(^2\)In order for \( \gamma_0^t \) to capture the macroeconomic conditions affecting the origination choices, we estimate the model for all banks reporting to the HMDA dataset but only use the \( \gamma_1^t \) for banks included in the final regressions. This implies including all the loan applications in the HMDA reporting in the estimations. The number varies between 17 million applications and 40 million application which constrains us to estimating the model as a linear probability model.

\(^3\)This measure is joint work with Lopez-Quiles and Petricek (2018).
3.5 IV Estimation of the Effect of Leverage on Risk Taking

In estimating the effect of leverage on risk taking we use the two instruments, explained in detail above. As argued before the instrumented deposits and the direct measure of exposure to local unemployment shocks are exogenous to risk taking. The instruments allow us to estimate two effects: (i.) the effect of the two instruments on leverage, and (ii.) the effect of leverage on risk taking.

More specifically we run the following two specifications:

$$\hat{\gamma}_{b,t} = \beta_0 + \beta_1 lev_{b,t} + \eta_b + \delta_t + \epsilon_{bt}$$  \hspace{1cm} (5)

Where $lev_{b,t}$ is the endogenous variable, measured as the leverage ratio, $\eta_b$ are bank fixed effects and $\delta_t$ are time fixed effects which are included to control for different baseline risk preferences of banks as well as a regulatory environment that may change over time. This equation is then estimated by IV, where the endogenous variable $lev_{b,t}$ is instrumented with one of the two instruments: either $\Delta dep_{bt}$ or $\Delta exp_{bt}$, depending on the model. It is also estimated by OLS, in order to compare the coefficients of interest.

The results of the estimations are presented and discussed in the next section.

4 Results

Table 4 presents the results of the estimations for both instruments. Column (1) shows the biased OLS estimate of regressing the risk taking measure on leverage ratio. Columns (2) and (3) present the results using the deposit supply shocks as an instrument, while columns (4) and (5) present the results employing the exposure to changes in unemployment. Columns (2) and (4) show the first stage of the two IV regressions, while Columns (3) and (5) show the second stage. Both sets of results are consistent in terms of sign, so we will focus on the latter in explaining them. All standard errors are clustered at the bank level.
First we find that the naive OLS approach severely underestimates the effect of leverage on risk taking behaviour, the result being close to zero and statistically insignificant. This bias may go some way towards explaining findings in the previous literature that leverage leads to less risk taking, or has no effect (see for example Altunbas et al. (2007) and Jacques and Nigro (1997)).

All our results provide evidence for a positive effect of leverage on banks’ risk taking due to limited liabilities.\footnote{Since leverage is measured as leverage ratio (core capital over total average assets), a negative sign implies that an increase in leverage (a decrease in leverage ratio) increases risk taking by banks.}

\begin{table}[h]
\centering
\caption{Risk Measure}
\begin{tabular}{lccccc}
\hline
 & (1) & (2) & (3) & (4) & (5) \\
\hline
Leverage & -0.0000902 & -0.0118** & -0.00564** & \\
 & (0.0000801) & (0.00557) & (0.00240) & \\
IV $\Delta$ Unemployment & 5.462**** & & & \\
 & (1.287) & & & \\
IV $\Delta$ Deposits & & -25.12**** & & \\
 & & (2.817) & & \\
Non-Current Loans & -0.163**** & -0.00255*** & -0.164**** & -0.00155**** \\
 & (0.00404) & (0.000913) & (0.00403) & (0.000400) \\
Constant & 0.0461**** & 9.776**** & 0.160*** & 10.35**** & 0.101**** \\
 & (0.000977) & (0.0290) & (0.0544) & (0.0713) & (0.0234) \\
Time FE & Yes & Yes & Yes & Yes & Yes \\
\hline
\end{tabular}

\textit{Stanard errors in parentheses}
\quad ** \ p < 0.05, *** \ p < 0.01, **** \ p < 0.001
\end{table}

We also investigate the possibility of nonlinear effects of leverage on risk taking by running our IV regression through 2SLS and interacting the predicted endogenous variable with dummy variables denoting different deciles of the leverage distribution. We find no significant pattern in the estimation and that most coefficient on these dummies are not statistically significant. The takeaway of this
analysis is that the effect of leverage in our data does not vary significantly along the distribution of leverage.

Quantitatively our results state that a one point increase in the leverage ratio\(^5\) generates an .0118 decrease in our risk taking measure. The mean of the risk measure in our data is .038, which implies that a one point increase in the leverage ratio produces a 31% decrease risk taking when compared to the average.

For a specific example, assume that two banks are identical except for their leverage ratios, which differ by one percentage point. Assume that they receive the same application for a mortgage loan with average Loan-to-Income. Our estimates suggest that this application has an expected probability of being originated in the bank with the higher leverage (i.e. lower leverage ratio) of \(\gamma_0 + \gamma_i LTI\) whereas the expected probability of origination for the less leveraged bank is \(\gamma_0 + (\gamma_i - .0118) LTI\). Evaluating these probabilities at the average of our estimates and at the average loan-to-income we obtain that the more leveraged bank has a 3.1% higher probability of originating the loan.

Note also that this wedge between the probabilities of acceptance increases with the loan-to-income ratio. Meaning that the higher the loan-to-income of the applicant the larger the difference in expected acceptance probabilities between the more and less leveraged bank.

Our results have policy relevant implications in terms of the aggregate level of risk in the banking system. The estimations show that more levered banks are more likely to take on riskier projects due to limited liability incentives which implies that curbing leverage has the added benefit of reducing banks’ risk taking.

\(^5\)Leverage ratio is defined as core capital divided by assets. Thus an increase in the leverage ratio is associated with a decrease in bank leverage.
thereby producing a more resilient banking system.

5 Robustness

In our main analysis we use as the outcome a risk measure based directly on banks’ lending decisions. The novelty of this measure lies in the fact that it captures the responsiveness of origination behaviour to a proxy of risk, namely the loan to income ratio. One could however be concerned that $L_tI$ being our only proxy for risk in our estimation the $\gamma$ parameter may be capturing correlations of $L_tI$ with unobserved loan characteristics. In order to address this potential problem redo the above analysis with a second measure for bank risk taking.

We resort to a further data source (Consumer Financial Protection Bureau) to obtain mortgage delinquency rates at the county level. This data covers 470 counties for the period 2009-2015. The sample is representative and covers approximately 85% of all mortgage lending in the HMDA data. We use the 90 days delinquency rate for closed-end, 1-4 family residential mortgages. The full list of variables and sources is in Table 1 in the Appendix.

The idea of this procedure is that we can observe the delinquency rate at the county level and we obtain a number of predictors with the same level of disaggregation. Once we have established a prediction model for our outcome we can predict the delinquency rate at the loan level by using more disaggregated data from the HMDA.

From the county data we build a 3 years ahead average delinquency rate at the county level, which will be used as the outcome of our prediction model. We obtain a number of explanatory variables from the American Community Survey, the Internal Revenue Service and the Bureau of Labour Statistics.

Since the goal is to be able to predict the delinquency rate at the loan level we run the model including two sets of explanatory variables: i) regressors that we
observe both at the county and at the loan level; 2) regressors we only observe at the county level.

In order to find the best prediction model we estimate approximately 1000 models including combinations of the predictors. For each model we split our sample keeping 75% of the counties to estimate the model and 25% as out of sample fit and we estimate it 10 times. We evaluate the models by the out of sample Mean Squared Prediction Error and pick the one with the lowest MSPE. All models are estimated with and without time fixed effects.

More formally we estimate

\[ DR_{i,t} = \beta_0 + \beta_1 X^1_{i,t} + \beta_2 X^2_{i,t} + \beta_3 X^1_{i,j,t} X^2_{i,t} + \delta_t + \epsilon_{i,t} \]  

(6)

Where \( i \) denotes the county and \( X^1 \) is a set of predictors that is aggregated at the county level from the loan level data, whereas \( X^2 \) is purely county level data. We include the interaction term between the loan aggregated and the county level regressors to account for the covariances between the two and eliminate potential bias when we move to the loan level prediction.

Once we have estimated this model we use the coefficient and produce the following loan level prediction

\[ DR_{i,j,t} = \hat{\beta}_0 + \hat{\beta}_1 X^1_{i,j,t} + \hat{\beta}_2 X^2_{i,t} + \hat{\beta}_3 X^1_{i,j,t} X^2_{i,t} + \hat{\delta}_t, \]  

(7)

where \( j \) denotes a loan. The specification of the prediction model is displayed in Table 2 in the Appendix.

This procedure gives us an expected probability of default for all the loans in our sample. In order to use this as an outcome in our final stage we take the median forecasted probability of default for each bank in every year. We estimate this model using the IV discussed in the methodology section. The results of this procedure are displayed in Table 5.
Table 5: Regression Table

|                      | (1)       | (2)       | (3)       | (4)       | (5)       |
|----------------------|-----------|-----------|-----------|-----------|-----------|
| **Leverage**         | Median PD | Median PD | Median PD | Median PD | Median PD |
|                      | -0.0208** | -1.174**  | -1.025**  |           |           |
|                      | (0.00861) | (0.434)   | (0.374)   |           |           |
| **IV Δ Unemployment**| 6.099**** | -15.22****|           |           |           |
|                      | (1.541)   | (3.510)   |           |           |           |
| **IV Δ Deposits**    | -0.125****| -0.122**  | -0.126****| -0.104**  |           |
|                      | (0.00527) | (0.0550)  | (0.00527) | (0.0475)  |           |
| **Non-Current Loans**| 2.592**** | 10.05**** | 14.26***  | 10.95**** | 12.74**** |
|                      | (0.0884)  | (0.0378)  | (4.416)   | (0.181)   | (3.802)   |
| **Constant**         | 2.592**** | 10.05**** | 14.26***  | 10.95**** | 12.74**** |
|                      | (0.0884)  | (0.0378)  | (4.416)   | (0.181)   | (3.802)   |
| **Time FE**          | Yes       | Yes       | Yes       | Yes       | Yes       |
| **N**                | 25050     | 25040     | 25040     | 25040     | 25040     |
| **R^2**              | 0.253     | 0.092     | 0.092     |           |           |
| **F(instr.excl.)**   | 37.05     | 44.04     |           |           |           |

Standard errors in parentheses
** p < 0.05, *** p < 0.01, **** p < 0.001

Using this measure for risk taking yields results that are in line with the ones discussed in the previous section. OLS underestimates the effect of leverage on risk taking for this measure as well. Using the IV procedure yields a negative and significant effect of leverage ratio on risk taking. This once again translates into a positive relationship between bank leverage and risk taking. Our results do not differ between the two IV procedures. Using this risk measure, we find that an exogenous one point increase in the leverage ratio of a bank will lead to roughly a one percentage point decrease in the median probability of default of loans issued by that bank.

6 Conclusions

This paper addresses the question of the causal effect of changes in leverage on banks’ risk taking behaviour. We do so by constructing two instruments to overcome the endogeneity problems resulting from the potential simultaneity and reverse causality between risk decisions and the deposit market conditions. We instrument exogenous changes in leverage by building two instruments: i) one based on the geographical area unemployment changes; ii) one based on the geo-
graphical area deposit supply changes. In both cases we aggregate them using the local deposit share of banks.

We then build a new measure of risk taking behaviour based on the responsiveness of origination decisions to a measure of risk of loan applications (loan-to-income). We compute this measure at the bank/year level and use it as our outcome.

Our empirical analysis suggests that exogenous increases in leverage incentivises banks to take on more risk, i.e. to originate loans with riskier characteristics. We also employ a second measure of risk, the median predicted probability of default of originated loans, and find that higher leverage leads to a higher probability of default of issued loans. These results are consistent with a limited liability and moral hazard story put forth by some of the theoretical literature. They are novel in the empirical literature on leverage and risk taking, as previous work has found no relation or a negative relationship between leverage and risk.

These results have relevant policy implications in that they suggest that any measure that would reduce banks’ leverage would also decrease incentives to invest in risky assets, thereby considerably reducing systemic risk.

References

Acosta Smith, J., Grill, M., and Lang, J. H. (2017). The Leverage Ratio, Risk-Taking and Bank Stability. SSRN Scholarly Paper ID 2990215, Social Science Research Network, Rochester, NY.

Altunbas, Y., Carbo, S., Gardener, E. P., and Molyneux, P. (2007). Examining the Relationships between Capital, Risk and Efficiency in European Banking. *European Financial Management*, 13(1):49–70.

Ashraf, B. N., Arshad, S., and Hu, Y. (2016). Capital Regulation and Bank Risk-Taking Behavior: Evidence from Pakistan. *International Journal of Financial Studies*, 4(3):16.
Bartik, T. J. (1993). Who Benefits from Local Job Growth: Migrants or the Original Residents? *Regional Studies*, 27(4):297–311.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654.

Dell’Ariccia, G., Igan, D., and Laeven, L. U. (2012). Credit booms and lending standards: Evidence from the subprime mortgage market. *Journal of Money, Credit and Banking*, 44(2-3):367–384.

Diamond, D. W. and Rajan, R. G. (2001). Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking. *Journal of Political Economy*, 109(2):287–327.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2017). Bartik instruments: What, when, why and how. *Unpublished Working Paper*.

Jacques, K. and Nigro, P. (1997). Risk-based capital, portfolio risk, and bank capital: A simultaneous equations approach. *Journal of Economics and Business*, 49(6):533–547.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

Koeln, M. and Santomero, A. M. (1980). Regulation of Bank Capital and Portfolio Risk. *The Journal of Finance*, 35(5):1235–1244.

Koudstaal, M. and van Wijnbergen, S. (2012). On Risk, Leverage and Banks: Do highly Leveraged Banks take on Excessive Risk? Tinbergen Institute Discussion Paper 12-022/2/DSF31, Tinbergen Institute.

Laeven, L. and Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics*, 93(2):259–275.
Lopez-Quiles, C. and Petricek, M. (2018). Deposit insurance and bank risk-taking.  
*ADEMU working paper series*, (2018/101).

Ohlrogge, M. (2017). Bank Capital and Risk Taking: A Loan Level Analysis.  
*SSRN Electronic Journal*.

Shriees, R. E. and Dahl, D. (1992). The relationship between risk and capital in commercial banks.  *
Journal of Banking & Finance*, 16(2):439–457.
## Appendix

### Variable list

| VARIABLE                          | SOURCE   | AVAILABILITY |
|-----------------------------------|----------|--------------|
| median ltv of originated loans for purchases or refinancing | HMDA     | loan         |
| median income of borrowers        | HMDA     | loan level   |
| median borrowed amount            | HMDA     | loan level   |
| median housing cost               | ACS      | county level |
| median housing cost to median income | ACS    | county level |
| median income                     | ACS      | county       |
| share households with a second mortgage | ACS    | county level |
| median value of a housing unit    | ACS      | county level |
| average monthly mortgage payment  | IRS      | county level |
| unemployment rate                 | BLS      | county level |

Data sources are:
HMDA - home mortgage disclosure act data
ACS - American Community Survey data
IRS - Internal Revenues Service
BLS - Bureau of Labor Statistics
# Delinquency rate prediction model

Table 2: Prediction Model

|                                      | (1)          |
|--------------------------------------|--------------|
|                                      | dr3c         |
| Median lti of origin. loans          | -0.482       |
|                                      | (0.686)      |
| Median income of borrowers           | -0.0621***   |
|                                      | (0.0145)     |
| Median household income              | 0.0000254    |
|                                      | (0.0000144)  |
| Med. housing cost to household inc.  | 315.1***     |
|                                      | (71.32)      |
| % of households with 2nd mortg.      | 0.0417       |
|                                      | (0.0343)     |
| Average mortgage payments            | 0.000217*    |
|                                      | (0.000110)   |
| Unemployment                         | 0.0000428*** |
|                                      | (0.00000523) |
| Interaction terms                    | YES          |
| State fixed effects                  | YES          |
| Constant                             | -1.223       |
|                                      | (1.485)      |
| N                                    | 1606         |
| $R^2$                                | 0.793        |
| Adj. $R^2$                           | 0.783        |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Figure 5: Median predicted probability of default; (average across banks in the final estimation sample)