BANK OF ENGLAND

# Staff Working Paper No. 912

# Organisational culture and bank risk

Joel Suss, David Bholat, Alex Gillespie and Tom Reader

March 2021

BANK OF ENGLAND

# Staff Working Paper No. 912

# Organisational culture and bank risk

Joel Suss,[1] David Bholat,[2] Alex Gillespie[3] and Tom Reader[4]

## Abstract

Existing research has largely relied on employee surveys to measure organisational culture despite the significant shortcomings of this approach. We use multiple, unobtrusive sources of data to gain rich insights into bank culture without ever having to ask employees to 'show us your culture'. Our measure is based on 20 individual indicators from six different sources, including information on internal fraud cases, customer complaints, and the quality of regulatory submissions. We use this data to investigate the hypothesised relationship between organisational culture and bank risk. We find robust evidence that poor culture leads to substantially higher risk, demonstrating the importance of bank culture for prudential outcomes.

**Key words:** Culture, bank risk, supervision.

**JEL classification:** G21, G30, L25, Z1.

(1) Bank of England. Email: joel.suss@bankofengland.co.uk
(2) Bank of England. Email: david.bholat@bankofengland.co.uk
(3) London School of Economics. Email: a.t.gillespie@lse.ac.uk
(4) London School of Economics. Email: t.w.reader@lse.ac.uk

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We are grateful to Marco Bardoscia, James Brookes, Marcus Buckmann, Chris Faint, Maxwell Green, Abubakr Karbhari, Kate Laffan, Alex Madewell, Paul Robinson, Misa Tanaka, Arthur Turrell, Eryk Walczak and an anonymous referee.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email enquiries@bankofengland.co.uk


ISSN 1749-9135 (on-line)

# Introduction

"As supervisors, we cannot go into a firm and say 'show us your culture'" – Andrew Bailey, 2016[1]

The organisational culture of banks is thought to be an important factor shaping their safety and soundness. Commentators often cite 'bad' or 'toxic' cultures as the root cause of major prudential and conduct failings in UK financial services in recent years, from the financial crisis of 2008 (Group of Thirty, 2018), to Libor rate-fixing (Salz and Collins, 2013), PPI mis-selling (Parliament, 2016), and rogue trading.[2]

However, while the link between poor organisational culture and bad banking outcomes is often discussed, there is surprisingly little research that has investigated this link empirically. This is predominantly due to measurement difficulties – identifying a bad banking culture before a crisis presents considerable challenges. As Andrew Bailey (2016) noted in a speech while Deputy Governor of the Bank of England and CEO of the Prudential Regulation Authority (PRA), banking supervisors cannot simply say to firms, 'show us your culture', and expect a satisfactory (or truthful) answer. Even so, employee self-reports remain the prevailing approach to measuring culture. Scholars and regulators typically rely on staff surveys and interviews with board members and senior executives to assess organisational culture (Gordon and DiTomaso, 1992; Graham et al., 2017; Guiso et al., 2015; Nuijts and Haan, 2013; O'Reilly III et al., 2014).

Yet, research in the behavioural sciences suggests that these sources are generally unreliable. There are a number of reasons for this. First, surveys and interviews lend themselves to impression management, i.e. deliberate attempts by those under scrutiny to create a favorable perception that may be at odds with reality (Antonsen, 2009; Bolino et al., 2008). Second, and less intentionally, employees embedded in particular organisations tend to take for granted their cultural context as 'normal' when, in fact, it may be anomalous and specific to that firm (Noort et al., 2016). This can make employees unreliable witnesses even if they are committed to revealing the truth about their firm's culture. Indeed, because culture shapes how people respond to surveys, it is not clear whether responses are perceptions or products of culture (McSweeney, 2002). Relatedly, employees who engage in unethical or risky acts rarely consider themselves to

[1] "Culture in Financial Services: A Regulator's Perspective", 9 May 2016.
[2] For example, the UBS rogue trader Kweku Adoboli, who lost the bank £1.74bn, has described the prevailing culture as a major contributing factor in statements to the BBC.

be doing so (Moore et al., 2012), with people justifying such acts (e.g., inflating results) as being helpful to an institution (Umphress et al., 2010). Finally, staff surveys and interviews tend to be non-representative of the entire population of staff – important swathes of an organisation, e.g. senior executives or employees who are disengaged, usually refrain from self-reporting, potentially biasing results (Heckman, 1979).

Given these problems with self-reported measures of organisational culture, a nascent literature has emerged which uses data gleaned *unobtrusively*, i.e. without explicitly asking firms to 'show us' their culture via employee self-reports (Reader et al., 2020). For example, recent studies have analysed online employee workplace reviews (Corritore et al., 2019; Moniz and Jong, 2014), internal email communications (Goldberg et al., 2016; Srivastava et al., 2018), annual report publications (Fiordelisi and Ricci, 2014; Gupta and Owusu, 2019; Nguyen et al., 2019), and other textual disclosures, e.g. company websites (Grennan, 2019a).

Although these studies represent advances in the assessment of organisational culture, they are not without limitations. Some data sources, like online employee workplace reviews, may be even more unrepresentative than staff surveys. Others, like the text in annual reports or on company websites, also clearly lend themselves to impression management. A consistent issue in these studies is that they typically rely exclusively on one source of unobtrusive data (e.g. earnings call transcripts; Li et al. (2019)) or a series of related sources (e.g. competing online employee review sites; Grennan (2019b)) to measure culture. Organisational culture, being distributed and pervasive, is unlikely to be captured by a single isolated measure. For example, detailed insight on the values of specific units or types of behaviour typically studied through surveys may be difficult to detect with data derived from a single source (e.g., earning call transcripts, or employee online reviews which only capture a small sample of an organisation's population). Instead, what is needed is a battery of measures probing different aspects and levels of the organisation.

This paper is a step in that direction, providing a richer picture of bank culture than has been available to date. It does so by exploiting a diverse array of unobtrusive sources of culture data, many of which are unique, that relate more clearly and exclusively to different dimensions of organisational culture. These include data quality metrics derived from regulatory return submissions; diversity metrics pertaining to banks' board and senior management; data on customer complaints and how they are handled by banks; whistleblowing referrals; data on internal fraud cases and costs; and capital requirement information. Moreover, whereas previous studies have often been limited to assessing culture at large institutions, our data covers the majority of UK banks and building societies, from multi-national universal institutions to regional building societies.

Altogether we define and measure 20 indicators of bank culture. Each of these were selected for conforming to one of the dimensions of culture identified in the Organizational Culture Profile (O'Reilly III et al., 2014, 1991), as well as an additional dimension – risk orientation – which is of particular relevance in the banking sector. For example, in order to assess a bank's customer orientation, we measure the proportion of complaints received which are closed slowly (defined as longer than 8 weeks). To measure a bank's ethical culture, we look at the number of internal fraud events as a proportion of the total operational risk events reported by the firm.

After describing the data, we test the hypothesis that poor organisational culture leads to bad prudential outcomes. To do so, we first create a summary culture score by aggregating the underlying indicators. We then examine whether culture relates to bank risk, using a standard (inverse) measure, the z-score distance to default metric. We find strong evidence of a link between organisational culture and bank risk – banks with poorer cultures are substantially more risky. Our findings are robust to different measures of risk, different constructions of the summary culture score, and different empirical specifications. To mitigate endogeneity concerns, we present results from instrumental variables and coarsened exact matching analyses. These likewise show a substantive link between poor organisational culture and bank risk.

The rest of this paper proceeds as follows: Section 2 surveys relevant existing literature and presents the guiding conceptual framework used for the construction of our culture measures; Section 3 details the data utilised; Section 4 provides the results and discussion; and, finally, Section 5 concludes with some suggestions for future research.

# 2 Unobtrusive Indicators of Culture

The term 'unobtrusive measure' was coined by Webb et al. (1966) to describe the value of using non-reactive methodologies – where data is collected and analysed without engaging participants. The key benefit of unobtrusive measures is that, compared to methodologies such as surveys where "the processes involved in measurement affect the value obtained for the variable" (Sechrest and Phillips, 1979, p .3), issues such as the social desirability in responses and observer effects can be addressed (Webb et al., 1966).

Drawing on this foundation, Reader et al. (2020) outline the construct of an 'unobtrusive indicator of culture' (UIC) and conduct a systematic review of the academic literature that use unobtrusive measures. A UIC refers to a single measure of organisational culture based on data collected without engaging employees. Measuring culture in this way addresses the social psychological finding that, just as attitudes do not necessarily correspond to behaviour (Wicker, 1969), the values and norms espoused by institutional members may not necessarily correspond to practices (Hill et al., 2014). Through drawing on naturally occurring data, which is often collated anonymously from across an institution (e.g., employee online reviews), extemporaneous (e.g., executives responding to questions during an earnings call), and revealing of values (e.g., institutional reward systems), analyses of culture are rooted in instantiations of organisational values rather than assessments. This, intuitively, seems more useful for capturing data on practices that are associated with poor outcomes. For instance, in an organisation where there are conditions that create risk (e.g., acceptance of unethical conduct, poor workforce management), gaining access to undertake a comprehensive survey of employees may be challenging, and reporting biases will likely influence the data collected. This is illustrated by research in domains such as organisational safety, where the paradox of assessing safety culture through surveys has long been recognised (Guldenmund, 2007; Mahler, 2009; Probst and Graso, 2013; Reason, 2000; Waring, 2005; Westrum, 2004), with the factors that lead to accidents (e.g., normalisation of unsafe behavior, poor management, blame, denying problems, lack of learning) potentially skewing the quality of data collected on safety values and practices from staff.

Culture measurement in financial services presumably has similar confounds, for instance in organisations with conduct problems, and thus may be particularly useful for detecting problematic bank cultures. As a result, UICs may have practical implications for bank regulators, among other stakeholders and market observers. Supervisors are often interested in monitoring organisational culture as a barometer of their safety and soundness, and so UICs are likely to be promising.

This paper leverages the conceptual framework articulated by Reader et al. (2020) and access to confidential regulatory data to develop a set of bank-specific UICs. Our search and selection of UICs was guided by two considerations. First, we collect data which is consistently measured over a considerable amount time for a wide range of banks, both big and small. Second, in order to guide our collection of data, we adopt a widely used schema of categorisation, the Organizational Culture Profile (OCP) (O'Reilly III et al., 2014). The OCP splits culture into six distinct dimensions: adaptability, inclusivity[3], integrity, results orientation, customer orientation, and detail orientation. We follow Grennan (2019a) by adding a seventh dimension, risk orientation, to conform more closely to the specific characteristics of the banking sector.

Adaptability refers to the willingness to innovate and the pace of internal change; integrity refers to firms with high ethical standards; inclusivity refers to an employee-oriented, inclusive and cooperative culture; results orientation indicates firms that focus on expectations and performance-related goals; customer orientation refers to the propensity to listen to customer needs and respond to their complaints; and detail orientation refers to an analytical and precise culture. The additional dimension risk orientation is related to the concept of 'risk culture' (Power et al., 2013) and can be thought of as the relative appetite for risk taking and the degree to which risk management is holistic and embedded throughout the organisation (Leaver and Reader, 2019).

# 3 Data

We gather culture data for 150 PRA-regulated banks and building societies (referred to hereafter simply as banks) on a quarterly basis between the years 2014-2020. Banks are either headquartered in the UK or subsidiaries of international firms. The UICs are derived from six sources: i) data quality metrics (DQMs) derived from regulatory return submissions; ii) diversity data from the Approved Persons database; iii) customer complaint reports; iv) whistleblowing referrals obtained from PRA intelligence; v) reports of internal fraud cases and costs; and vi) balance sheet and capital requirement information. From these sources, we compute a total of 20 indicators grouped into five of the seven above-listed dimensions of culture. We describe these variables below.

[3] We have re-labelled the original term 'collaborative' as inclusivity to reflect the growing understanding of what collaboration presupposes and involves.

### *3.1 Data quality metrics as indicators of detail orientation*

Banks are required to submit regulatory returns to the PRA on a regular basis. The promptness and quality of these returns varies, and so we can compute and track a host of data quality metrics (DQMs). Specifically, we calculate five DQMs: i) the number of submission validation failures;[4] ii) the number of individual returns submitted past the required deadline; iii) the number of days submitted late per submission; iv) the number of days a firm submits early; and v) the number of plausibility flags raised by Bank of England analysts.[5] Each of the indicators related to late or early submissions are standardised by the total number of regulatory return modules (i.e. top-level returns such as COREP 001.a), while validation failures and plausibility flags by the number of templates (i.e. tables within modules) submitted in that quarter. This accounts for the impact a bank's size or breadth of business model has on the raw figures and allows for relative comparison.

All the DQMs fall within the *detail orientation* dimension – they provide insight into the extent to which banks emphasise attention to detail and precision.[6] Banks which regularly report their data on time or early to the regulator, and which have a relatively low number of plausibility flags and validation failures, exhibit strong detail orientation.

### *3.2 Approved Persons data for measures of inclusivity*

The Financial Services and Markets Act 2000 requires that individuals in the most senior roles at banks seek approval from the regulator prior to assuming their roles. For an individual to be approved, the regulator must be satisfied in the individual's integrity, competence, and financial soundness. The details of which roles require vetting has changed over time, most recently with the advent of the Senior Managers Regime (from March 2016).

Information about individuals in senior leadership roles is captured in an 'Approved Persons' database, providing information on the gender, age and tenure of all individuals subject

---

[4] These are reporting rules set by the European Banking Authority which can result in the rejection of the submitted file or a warning without rejection.
[5] These are automatic red flags raised, for example, in response to a large change from one quarter to another in a reported figure, that instigates a manual verification and, in cases where it is deemed material, the reporting firm is notified and required to take corrective action.
[6] The quality of reporting also has direct financial implications for firms. In 2020 Citibank UK was fined £44mn by the PRA for failures in regulatory reporting governance and controls.

to approval.[7] As a result, we are able to capture quarterly snapshots of four UICs: i) the overall proportion of women among all approved persons at the firm; ii) the difference in the proportion of women between non-executive directors and executives; iii) the standard deviation of age; and iv) the difference in average age between the non-executive and executive directors. We suppose that there is a concave relationship between i), iii) and iv) and inclusivity. Rather than the simple proportion of women, we instead take the Blau index. This ranges from 0 to 0.50, with 0.10 and 0.90 being treated equivalently. We also bin the gender difference by deciles so that -0.50 and 0.50 are in the same bin. For the age difference variable we similarly bin at intervals of five years.

These indicators of the diversity of the people approved to carry out controlled and senior manager functions relate to the inclusivity dimension of the OCP. We consider banks where there is a diversity of genders and ages represented at senior levels as more inclusive and collaborative, and collaboration is expected to lead to better decision-making (Salas et al., 2020). Moreover, a bank that exhibits a greater disconnect between the board and executive in terms of gender and age composition are considered to lack inclusivity.

### *3.3 Complaints reports for indicators of customer orientation*

Every UK-regulated bank must report the number of customer complaints it receives to the UK's Financial Conduct Authority on a half-yearly basis. From these reports, which we convert to a quarterly-basis through interpolation, we calculate five UICs: i) the proportion of complaints outstanding at the end of the period; ii) the proportion of complaints upheld by the bank; iii) the proportion of complaints that are closed slowly, i.e. takes the bank longer than 8 weeks to close; iv) the amount of redress paid per closed complaint; and v) the number of complaints held per £bn in balance sheet assets.

These indicators provide us with information on the customer orientation dimension of culture and, as with customer complaints data in other industries (Reader and Gillespie, 2020), reveal two issues: a failure in service provision, and problems in dealing with that failure. Banks that have a large volume of complaints, a high proportion of complaints upheld, or a large amount of redress paid per complaint, are serving customers relatively poorly. Measures that

[7] See also related work by Suss et al. (in press) which describes this data in greater detail and explores how gender and age diversity have evolved over time in UK banks.

give an indication of how a bank responds to customer complaints, i.e. the proportion of complaints outstanding or closed slowly, similarly provide an indication of a bank's customer orientation.

### *3.4 Whistleblowing and internal fraud data as a measure of integrity*

To measure a firm's integrity, we combine data from two sources: whistleblowing referrals and regulatory returns on operational risk. From these sources we compute three UICs: i) the number of whistleblowing referrals received from the firm and followed-up by PRA supervisors, ii) the proportion of operational risk events which are due to internal fraud, and iii) the proportion of monetary loss due to internal fraud events.

Banks with a relatively high incidence of internal fraud and costs incurred because of it can be considered to lack integrity. Similarly for whistleblowing. Whistleblowing involves an allegation by an employee that they have observed activity in a firm that is illegal or otherwise contrary to the public interest.[8] A high number of whistleblowing referrals to the regulator implies a lack of integrity within the firm.

### *3.5 Balance sheet and capital requirements data for measures of risk orientation*

A core aspect of prudential regulation is monitoring bank balance sheets and ensuring firms adhere to capital requirements. A bank's position relative to a relevant peer group (defined as the other banks in a given firm's regulatory division at the Bank of England) provides insight into its relative risk orientation. For UICs, we therefore take the difference between a firm's average risk-weight and the average risk-weight of its peer group. A positive value likely indicates that the bank has a relatively higher risk appetite and vice versa. Similarly, we compute the difference between a bank's capital buffer ratio (defined as the difference between regulatory capital and capital requirements divided by total risk-weighted assets) and the average buffer ratio of its peer group. A positive value (a relatively higher buffer ratio) in principle means the bank is safer, from which we might infer a lower risk orientation, and vice-versa. Finally, we use the ratio between a firm's Pillar 2A capital requirement (a capital add-on that is applied at the discretion of the PRA) and its total capital requirement as an indicator of risk orientation. A bank

---

[8] See details on formal whistleblowing on the Bank of England website.

that has a relatively high proportion of Pillar 2A capital requirements might in part be a reflection of its risk culture.

All the UICs and the dimensions they represent are summarised in Table A.1 in the Annex.

### *3.6 Aggregating UICs into dimensions and a summary score*

In isolation, an individual UIC is unlikely to be sufficient in representing a bank's culture. However, by aggregating together UICs, we expect a clearer picture of culture to emerge. We therefore aggregate the UICs into dimensions as per the OCP categorisation described above, as well as altogether into a summary culture score. Our empirical analysis in the next section takes these aggregate scores as the explanatory variables of interest.

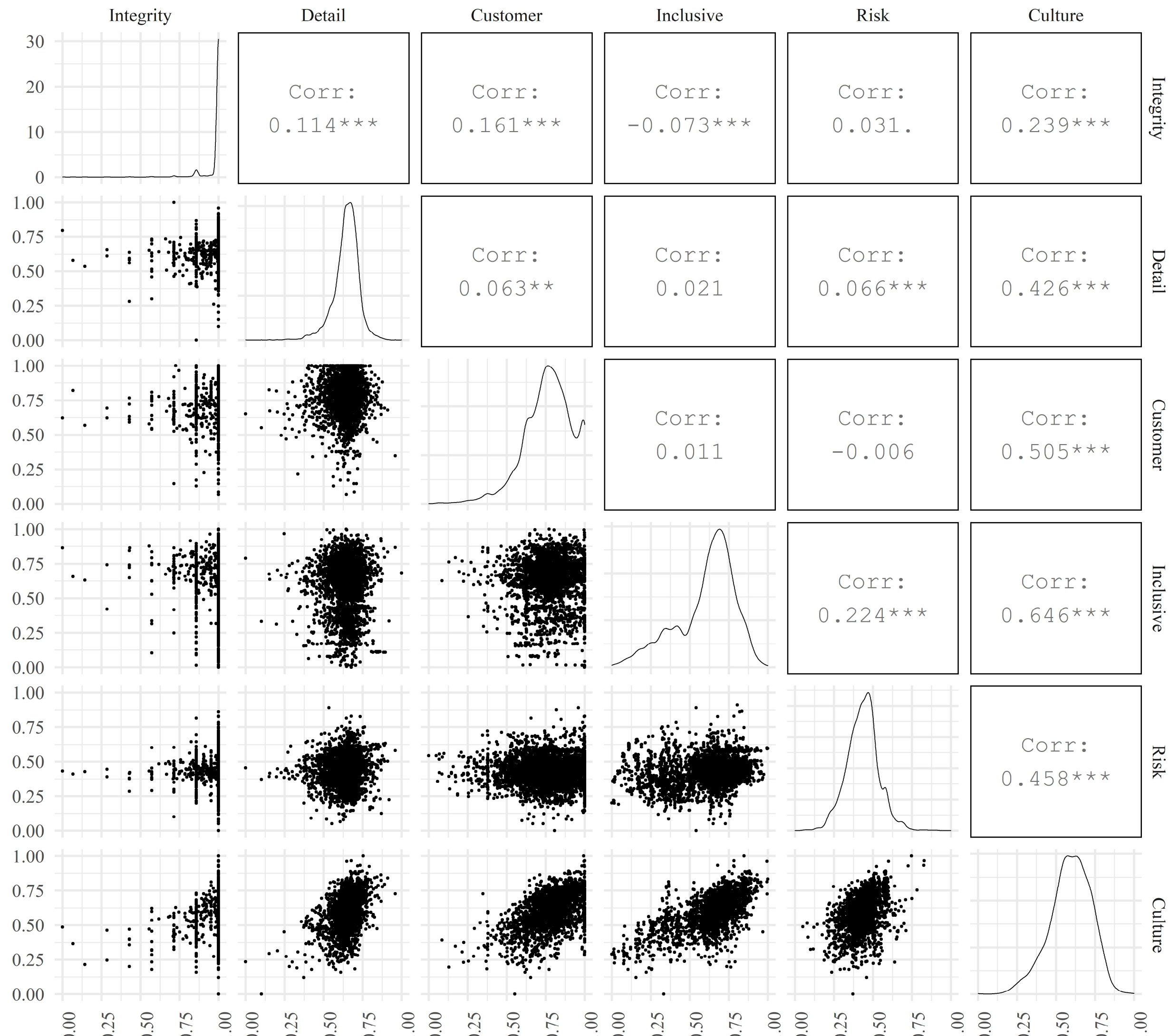
The aggregation is done as follows: we scale each UIC to range between 0 and 1, preserving the original distribution. We then take a simple average of all UICs, reversing the sign of those that are considered a priori to be symptomatic of a poor culture, i.e. all those other than the average days reporting early, Blau index of the proportion female, standard deviation of age, and the relative difference in capital buffer. As such, higher scores indicate healthier cultures and lower scores unhealthy ones. We handle missing values as follows: for each of the dimensions, we take the average of the relevant UICs as long as 50% or more of the UICs are non-missing. If 50% or more of the UICs underlying one or more of the dimensions are missing, the aggregate score is considered to be missing as well.[9] In the Annex (Table A.2) we investigate the robustness of our results to alternative ways of computing aggregate culture scores, including in relation to treating missing values – in particular we take the average for complete cases only, as well as the first principal component. We also define the summary score as the average of all the constituent dimensions rather than UICs, and iteratively remove one of the five dimensions from the summary culture score in turn.

Figure 1 provides a correlation matrix and distribution for each of the aggregated scores. We can see from the matrix that the values of the correlation coefficients between dimension scores are low, with 7 out of 10 being less than 0.1 in absolute terms and a maximum of 0.224. This suggests that the dimension aggregates are indeed capturing different facets of

[9] We allow for an exception to this approach for the integrity dimension, which is that if all three UICs are missing we force to aggregate score to missing. This is because we have far more missing data for the internal fraud UICs, therefore we allow a score to be computed if only 1 out of the 3 UICs for the integrity dimension are present.

organisational culture. On the other hand, the overall culture score correlates moderately to strongly with all dimensions, with a maximum of 0.646, giving credence to simple averaging as a means to adequately summarise the underlying cultural dimensions.

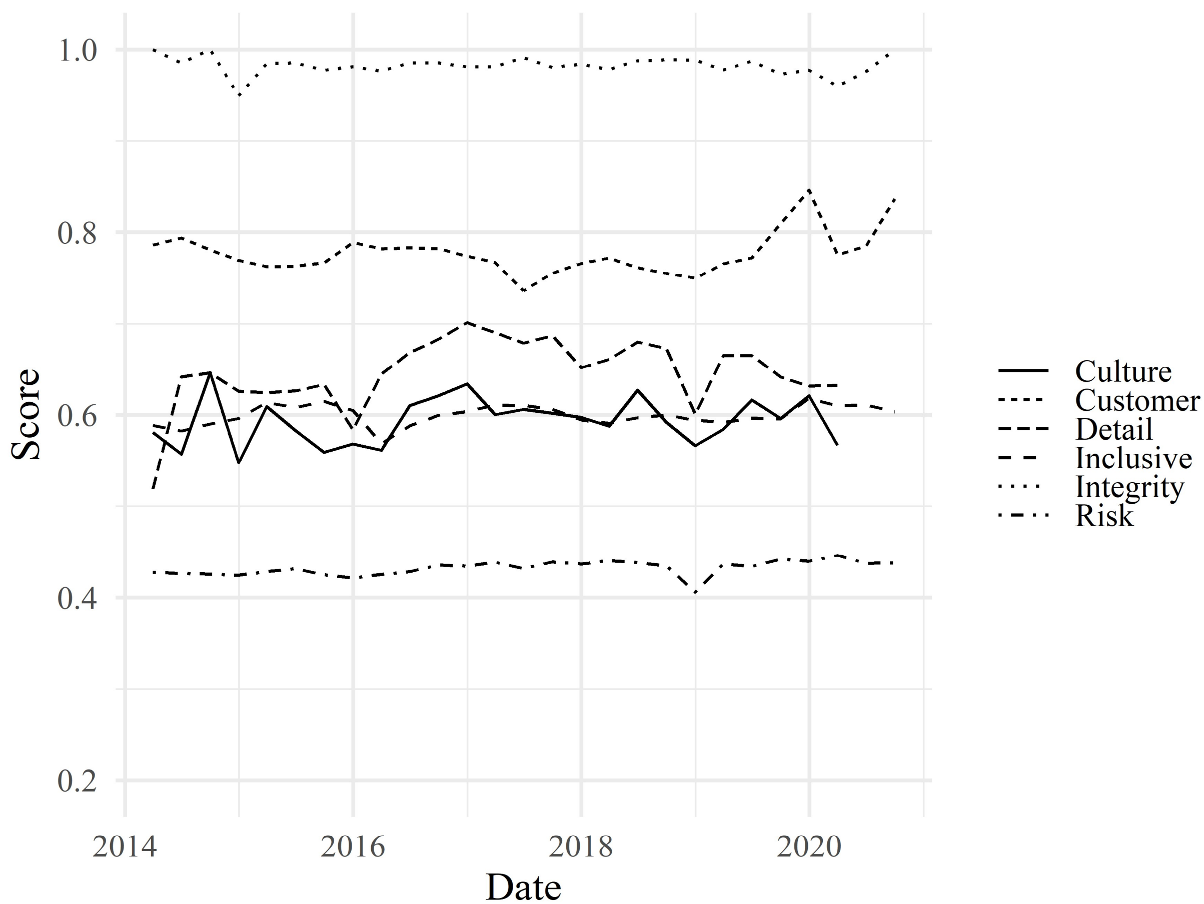
**Figure 1: Correlation matrix and distributions for each culture score**




*Note: Higher values indicate healthier cultures. All aggregated scores are scaled to range between 0 and 1. The culture score is the average of each underlying UIC, whereas the other scores are averages of UICs pertaining to that dimension. All dimension measures are subject to there being less than 50% missing for any firm-quarter observation, whereas the summary culture score is missing if one or more of the dimensions is missing more than 50% (except for integrity, where all three UICs have to be missing).*

Figure 2 shows the trend of average scores by quarter over the period between 2014 and 2020. We can see that both the overall culture score and its constituent dimensions are largely static, with the exception of customer orientation, where the average increases slightly to above 0.8 in 2020, and inclusivity, which has also trended slightly upwards over the period. We also see that the quarterly average for risk orientation is the lowest out of any of the dimensions, hovering just above 0.4, while firms are on average high on integrity.

**Figure 2: Time trend of cultural scores, 2014-2020**




### *3.7 Bank risk measures*

We use a standard (inverse) measure of bank risk as our main outcome variable: the z-score distance to default measure. The z-score is calculated as:

$$Z = \frac{ROA + (Capital/Assets)}{\sigma_{ROA}}$$

where $ROA$ is a bank's return on assets (i.e. net profit divided by average assets), and $(Capital/Assets)$, otherwise known as the leverage ratio, is a bank's equity capital relative divided by total assets, and $\sigma_{ROA}$ is the standard deviation of ROA calculated using an 8-quarter window. The z-score calculates the number of standard deviations a bank's ROA has to drop by to offset its total regulatory capital. The higher the value of $Z$, the further away from default and thus the less risky the bank is. In our empirical analysis below, we use the log of the z-score.

In addition, we use a number of alternative outcome measures of bank risk: i) subjective assessments by PRA supervisors – these are on a scale from 1 to 4, with 4 being the highest risk notch[10]; ii) the Sharpe ratio – defined as return on equity (ROE) divided by volatility of ROE (using an 8 quarter window); and iii) the ratio of non-performing loans (loans 90+ days past due) divided by CET1 capital. We also vary the size of the window used to calculate the volatility of returns to 12 and 16 quarters.

### *3.8 Control variables*

We include a measure of bank size (natural log of total assets) as a control variable. We also include two tenure variables derived from the Approved Persons database: the average tenure in the firm in quarters for both the CEO and non-executive directors as a group. This is to account for the possibility that culture emanates from the top (O'Reilly III et al., 2014), and so banks with lower average tenure for these variables may have different cultures than those who with higher average tenure.

To account for differences in business model and balance sheet composition, we include three measures: i) the ratio of retail deposits to assets – known as the core deposit ratio; ii) the average risk weight (i.e. total risk-weighted assets divided by total assets); and iii) and the ratio of loans to assets. Finally, to account for a possible relationship between geographical dispersion of business activities, which has been shown to be related to bank risk (Berger et al., 2017), and culture, we include a ratio of foreign assets (defined as those outside the UK) to total assets. Table 1 provides summary statistics for all the UICs, outcome and control variables.

---

[10] See here for more details on the PRA's approach to supervision.

**Table 1: Descriptive statistics**

| Variable | N | Mean | St.Dev | Min | Median | Max |
|---|---|---|---|---|---|---|
| **Customer orientation** | | | | | | |
| Complaints Outstanding | 1953 | 12.87 | 16.43 | 0.00 | 7.61 | 100.00 |
| Complaints Upheld | 2086 | 33.94 | 25.68 | 0.00 | 32.90 | 100.00 |
| Complaints Closed Slow | 2080 | 5.72 | 12.41 | 0.00 | 0.58 | 100.00 |
| Redress Paid Per Complaint | 2066 | 220.01 | 591.75 | 0.00 | 35.85 | 6720.17 |
| Complaints Opened / Assets | 2079 | 109.40 | 248.37 | 0.00 | 32.89 | 2263.70 |
| **Detail orientation** | | | | | | |
| Validation Failures | 2101 | 0.38 | 0.33 | 0.00 | 0.30 | 2.50 |
| Plausibility Flags | 1652 | 0.08 | 0.12 | 0.00 | 0.04 | 0.95 |
| Lateness | 2101 | 0.16 | 1.31 | 0.00 | 0.00 | 23.53 |
| Late Days | 2076 | 0.09 | 0.52 | 0.00 | 0.00 | 8.48 |
| Early Days | 2101 | 3.64 | 4.30 | 0.00 | 2.43 | 36.75 |
| **Inclusivity** | | | | | | |
| Gender Blau | 2101 | 0.28 | 0.14 | 0.00 | 0.30 | 0.50 |
| Gender Difference | 1638 | 2.95 | 1.69 | 1.00 | 3.00 | 10.00 |
| Age S.D. | 2101 | 7.89 | 2.30 | 1.34 | 7.87 | 16.27 |
| Age Difference | 2100 | 2.44 | 1.08 | 1.00 | 2.00 | 6.00 |
| **Integrity** | | | | | | |
| Whistleblowing Referrals | 2025 | 0.17 | 0.63 | 0.00 | 0.00 | 7.00 |
| Operational loss: internal fraud | 411 | 0.49 | 2.26 | 0.00 | 0.00 | 21.45 |
| Operational events: internal fraud | 411 | 0.19 | 0.55 | 0.00 | 0.00 | 7.14 |
| **Risk orientation** | | | | | | |
| Avg RW Difference | 2082 | -0.39 | 16.32 | -49.75 | 0.30 | 57.78 |
| Capital Buffer Difference | 2101 | -5.51 | 19.79 | -117.03 | -3.39 | 220.80 |
| Pillar 2 / ICG | 2101 | 28.99 | 14.36 | 0.00 | 29.20 | 87.54 |
| **Controls** | | | | | | |
| Size | 2101 | 7.70 | 2.39 | 2.58 | 7.04 | 14.64 |
| Tenure Firm CEO | 2033 | 25.31 | 20.19 | 0.00 | 19.00 | 73.00 |
| Tenure Firm NEDs | 2101 | 16.78 | 9.73 | 0.00 | 15.50 | 68.00 |
| Total Deposits / Assets | 2014 | 66.97 | 28.81 | 0.00 | 79.32 | 94.81 |
| Avg Risk Weight | 2067 | 44.85 | 20.31 | 9.54 | 37.77 | 117.97 |
| Loans / Assets | 2008 | 54.81 | 29.73 | 0.00 | 68.69 | 92.13 |
| International Asset Ratio | 1987 | 23.18 | 29.78 | 0.00 | 4.65 | 100.00 |
| **Outcome measures** | | | | | | |
| Z-score | 1746 | 4.58 | 0.94 | 0.82 | 4.70 | 7.59 |
| SD(ROA) | 1741 | 0.18 | 0.28 | 0.00 | 0.08 | 2.25 |
| ROA | 1952 | 0.32 | 1.07 | -4.84 | 0.27 | 5.43 |
| Leverage ratio | 2100 | 10.45 | 8.55 | 1.41 | 7.77 | 82.38 |
| Sharpe ratio | 1715 | 4.10 | 3.48 | -4.30 | 4.27 | 12.89 |
| Non-Performing Loans / CET1 | 1508 | 20.28 | 26.45 | 0.00 | 11.21 | 179.38 |

*Note: Descriptive statistics for the years 2014-2020. Total firms = 150.*

# 4 Empirical Results

## *4.1 Univariate analysis*

We first examine the unconditional mean differences for observations that are above or below the global median z-score value. Table 2 shows that our summary bank culture score is significantly higher for those above versus below the median. This holds for the differences in each of the five cultural dimensions. Table 2 also shows mean differences and significance for each of the underlying UICs. These generally conform to expectations in terms of direction and significance.

**Table 2: Mean differences by variable, above or below median z-score**

| Variable | Above | Below | T Statistic | Difference |
|---|---|---|---|---|
| **Aggregate culture scores** | | | | |
| Integrity | 0.98 | 0.97 | 3.34 | 0.01*** |
| Detail | 0.67 | 0.64 | 6.87 | 0.02*** |
| Customer | 0.80 | 0.74 | 8.81 | 0.06*** |
| Inclusive | 0.65 | 0.62 | 3.31 | 0.03*** |
| Risk | 0.45 | 0.43 | 5.77 | 0.02*** |
| Culture score | 0.63 | 0.56 | 10.10 | 0.06*** |
| **Customer orientation** | | | | |
| Complaints Outstanding | 13.64 | 12.83 | 1.00 | 0.81 |
| Complaints Upheld | 28.74 | 38.18 | -7.71 | -9.44*** |
| Complaints Closed Slow | 5.34 | 6.56 | -2.00 | -1.22** |
| Redress Paid Per Complaint | 122.70 | 355.49 | -7.86 | -232.79*** |
| Complaints Opened / Assets | 95.83 | 128.55 | -2.71 | -32.72*** |
| **Detail orientation** | | | | |
| Validation Failures | 0.37 | 0.38 | -0.93 | -0.01 |
| Plausibility Flags | 0.06 | 0.08 | -4.59 | -0.03*** |
| Lateness | 0.04 | 0.26 | -3.57 | -0.21*** |
| Late Days | 0.08 | 0.11 | -1.36 | -0.04 |
| Early Days | 4.18 | 3.37 | 3.96 | 0.81*** |
| **Inclusivity** | | | | |
| Gender Blau | 0.30 | 0.27 | 3.51 | 0.02*** |
| Gender Difference | 3.02 | 2.81 | 2.27 | 0.21** |
| Age S.D. | 7.58 | 8.12 | -4.82 | -0.54*** |
| Age Difference | 2.32 | 2.51 | -3.61 | -0.19*** |
| **Integrity** | | | | |
| Whistleblowing Referrals | 0.13 | 0.25 | -3.78 | -0.12*** |
| Operational loss: internal fraud | 0.36 | 0.62 | -1.06 | -0.26 |
| Operational events: internal fraud | 0.21 | 0.22 | -0.12 | -0.01 |
| **Risk orientation** | | | | |
| Avg RW Difference | 1.17 | -0.03 | 1.63 | 1.2 |

| | | | | |
|---|---|---|---|---|
| Capital Buffer Difference | -3.83 | -6.89 | 3.49 | 3.07*** |
| Pillar 2 / ICG | 26.24 | 31.49 | -7.63 | -5.25*** |
| **Controls** | | | | |
| Size | 7.24 | 8.00 | -6.66 | -0.76*** |
| Tenure Firm CEO | 28.40 | 23.61 | 4.88 | 4.8*** |
| Tenure Firm NEDs | 18.71 | 16.32 | 5.20 | 2.39*** |
| Total Deposits / Assets | 65.46 | 71.28 | -4.28 | -5.82*** |
| Avg Risk Weight | 43.03 | 47.38 | -4.49 | -4.35*** |
| Loans / Assets | 56.84 | 55.27 | 1.09 | 1.57 |
| International Asset Ratio | 22.01 | 25.05 | -2.04 | -3.04** |

*Note: *p<0.1; **p<0.05; ***p<0.01. All aggregate culture scores are scaled to range from 0 to 1.*

## *4.2 Main results*

We now turn to the main empirical results. We specify a linear regression model with quarter fixed effects and robust standard errors at the bank level.[11] In order to mitigate the possibility of reverse causality, we lag all the independent variables by one quarter.

Table 3 provides the main results, with all variables mean-centred and scaled by their respective standard deviation. Column 1 of Table 3 provides the headline results with the summary culture score as the explanatory variable. We find that culture is statistically significant with a standardised coefficient of 0.242 Column 2 instead includes each of the dimensions together. Looking at Column 2 of Table 3, we see that the coefficients on the detail, customer, and risk orientation measures are significant, with standardised coefficients of 0.139, 0.136, and 0.251 respectively.

Columns 3-5 run the same regression as in Column 1 but for a subset of the data. Column 3 only includes large banks – defined as those with £50bn or more in total assets, Column 4 includes only mid-sized firms – between £1bn and £50bn in total assets, and Column 5 banks with less than £1bn in total assets. Here we see that the coefficient on the culture score remains significant regardless of the subset and ranges from 0.181 to 0.265

The final column of Table 3 collapses the data by time. The coefficient on the culture score remains significant at 0.242.

[11] While our panel is fairly long, we don't include firm fixed effects for two reasons. First, culture is known to change slowly in firms, and so within firm differences from the mean can be attributed to the noisiness of the underlying UICs rather than meaningful change in culture. Furthermore, a simple auto-regressive model reveals the scores to be highly persistent over time, with a coefficient on the lag term of 0.782, and so firm fixed effects are not recommended (Berger et al., 2017; Zhou, 2001).

**Table 3: Main regression results**

| | *Dependent variable:* Z-score | | | | | |
|---|---|---|---|---|---|---|
| | Full sample | Full sample | Large size (total assets > £50bn) | Medium size (£50bn > total assets >= £1bn) | Small size (£1bn > total assets) | Analysis of averages |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Culture | $0.242^{***}$ | | $0.181^{**}$ | $0.265^{**}$ | $0.201^{**}$ | $0.242^{**}$ |
| | (0.073) | | (0.083) | (0.117) | (0.097) | (0.104) |
| Integrity | | -0.026 | | | | |
| | | (0.033) | | | | |
| Detail orientation | | $0.139^{**}$ | | | | |
| | | (0.058) | | | | |
| Customer orientation | | $0.136^{**}$ | | | | |
| | | (0.057) | | | | |
| Inclusivity | | 0.115 | | | | |
| | | (0.090) | | | | |
| Risk orientation | | $0.251^{**}$ | | | | |
| | | (0.102) | | | | |
| Size | $-0.190^{**}$ | $-0.152^{*}$ | -0.030 | -0.091 | -0.103 | -0.065 |
| | (0.077) | (0.091) | (0.173) | (0.107) | (0.140) | (0.106) |
| Tenure CEO | $0.092^{**}$ | $0.082^{*}$ | 0.082 | 0.118 | 0.060 | 0.109 |
| | (0.046) | (0.043) | (0.103) | (0.073) | (0.064) | (0.090) |
| Tenure NEDs | 0.065 | 0.051 | 0.078 | 0.022 | $0.157^{*}$ | -0.008 |
| | (0.075) | (0.073) | (0.105) | (0.103) | (0.081) | (0.089) |
| Core deposit ratio | $-0.236^{***}$ | $-0.235^{***}$ | $-0.702^{**}$ | -0.085 | $-0.231^{**}$ | -0.062 |
| | (0.081) | (0.080) | (0.299) | (0.125) | (0.109) | (0.111) |
| Average risk-weight | $-0.140^{*}$ | -0.057 | $0.410^{**}$ | -0.012 | -0.182 | 0.089 |
| | (0.072) | (0.082) | (0.173) | (0.111) | (0.118) | (0.119) |
| Loans / assets | $0.151^{*}$ | $0.172^{**}$ | 0.391 | $0.206^{**}$ | 0.090 | -0.021 |
| | (0.077) | (0.074) | (0.374) | (0.090) | (0.126) | (0.124) |
| International ratio | $0.161^{*}$ | $0.185^{**}$ | -0.043 | $0.273^{**}$ | 0.041 | $0.308^{**}$ |
| | (0.088) | (0.087) | (0.173) | (0.136) | (0.113) | (0.149) |
| Time fixed effects | Yes | Yes | Yes | Yes | Yes | No |
| Clustering level | Bank | Bank | Bank | Bank | Bank | None |
| Number of banks | 121 | 121 | 12 | 57 | 63 | 121 |
| Observations | 1,532 | 1,532 | 154 | 684 | 694 | 121 |
| $R^2$ | 0.145 | 0.168 | 0.503 | 0.142 | 0.137 | 0.292 |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

All variables are mean-centred and scaled by their respective standard deviation. All independent variables are lagged by 1 quarter.

### *4.3 Alternative risk measures*

Next, we examine whether the findings seen in Table 3 are robust to alternative risk measures. Table 4 provides these results. Column 1 uses as a risk measure the subjective assessment of PRA supervisors, known as the PIF score (scale from 1-4, with 4 being the highest risk). The dependent variable for Columns 2 and 3 are two commonly used measures of bank risk – the Sharpe ratio and the ratio of non-performing loans to CET1 capital, respectively. The dependent variable in Columns 4 and 5 is again the z-score, but with 12 and 16 quarter windows used to calculate the standard deviation of ROA. Column 6 is the unlogged z-score. Finally, Columns 7 and 8 have as dependent variables the logarithmised z-score measure and Sharpe ratio but all independent variables are lagged by 8 quarters rather than 1 due to address potential endogeneity concerns arising because a component of the dependent measure (the standard deviation of ROA or ROE) is partially determined before the culture score (Berger et al., 2017).

Panel A of Table 4 shows the results for the summary culture score. The coefficient remains significant and in the expected direction for each of the alternative risk measures. Once again, the coefficients are substantive in size, with a one standard deviation increase in culture associated with a decrease in risk, however defined, by between 0.159 and 0.258 of a standard deviation (leaving aside the unlogged z-score, whose distribution exhibits a pronounced rightward skew).

Panel B of Table 4 provides the individual dimensions of culture. Across the different models, only the customer orientation coefficient is consistently significant, albeit not for the specification in Column 2. The coefficient on detail orientation is weakly significant in 5 out of the 8 models, and the coefficient on inclusivity is significant when the dependent variable is the PIF score, Sharpe ratio, as well as the models with 9 quarter lags of the independent variables. Integrity and risk orientation are significant in some models but are mostly insignificant. This inconsistent picture suggests that individual dimensions, with the exception of customer orientation, are not robustly related to risk holding the level of the other dimensions constant. On the other hand, the relationship between bank risk and the aggregate culture score, which effectively synthesises the individual dimensions into an overall picture of culture, appears robust.

**Table 4: Regression results for alternative bank risk measures**

| | *Panel A* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | PIF | Sharpe | Non-performing loans / CET1 | Z-score (12Q window) | Z-score (16Q window) | Exp(Z-score) | Z-score (8Q lags) | Sharpe (8Q lags) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Culture | -0.256$^{***}$ | 0.227$^{***}$ | -0.159$^{***}$ | 0.245$^{***}$ | 0.234$^{***}$ | 0.095$^{**}$ | 0.258$^{***}$ | 0.208$^{***}$ |
| | (0.054) | (0.052) | (0.054) | (0.081) | (0.082) | (0.046) | (0.069) | (0.060) |
| Time Fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clustering level | Banks | Banks | Banks | Banks | Banks | Banks | Banks | Banks |
| Number of banks | 146 | 120 | 127 | 121 | 120 | 121 | 114 | 114 |
| Bank controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,804 | 1,502 | 1,275 | 1,480 | 1,422 | 1,532 | 1,020 | 1,005 |
| $R^2$ | 0.279 | 0.204 | 0.233 | 0.151 | 0.153 | 0.091 | 0.195 | 0.144 |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | | | | | | | |
| | All variables are mean-centred and scaled by their respective standard deviation. All independent variables are lagged by 1 quarter, except for columns 7 and 8, where they are lagged by 8 quarters. | | | | | | | |

| | *Panel B* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | PIF | Sharpe | Non-performing loans / CET1 | Z-score (12Q window) | Z-score (16Q window) | Exp(Z-score) | Z-score (8Q lags) | Sharpe (8Q lags) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Integrity | -0.113$^{***}$ | 0.067$^{**}$ | -0.058 | 0.005 | 0.017 | -0.007 | 0.014 | 0.022 |
| | (0.035) | (0.029) | (0.040) | (0.033) | (0.034) | (0.022) | (0.034) | (0.031) |
| Detail orientation | -0.081$^{*}$ | 0.056 | -0.128$^{***}$ | 0.094$^{*}$ | 0.103$^{*}$ | 0.048 | 0.146$^{**}$ | 0.020 |
| | (0.042) | (0.036) | (0.045) | (0.054) | (0.056) | (0.033) | (0.065) | (0.048) |
| Customer orientation | -0.131$^{***}$ | 0.104$^{*}$ | -0.094$^{*}$ | 0.156$^{**}$ | 0.140$^{**}$ | 0.077$^{**}$ | 0.141$^{**}$ | 0.143$^{**}$ |
| | (0.050) | (0.056) | (0.051) | (0.065) | (0.065) | (0.035) | (0.062) | (0.064) |
| Inclusivity | -0.108$^{**}$ | 0.183$^{***}$ | -0.075 | 0.160 | 0.163 | 0.012 | 0.165$^{**}$ | 0.169$^{**}$ |
| | (0.051) | (0.060) | (0.053) | (0.102) | (0.101) | (0.054) | (0.070) | (0.069) |
| Risk orientation | -0.201$^{**}$ | 0.133$^{*}$ | -0.085 | 0.188$^{*}$ | 0.154 | 0.113$^{**}$ | 0.160 | 0.097 |
| | (0.089) | (0.074) | (0.093) | (0.109) | (0.108) | (0.045) | (0.117) | (0.092) |
| Time Fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clustering level | Banks | Banks | Banks | Banks | Banks | Banks | Banks | Banks |
| Number of banks | 146 | 120 | 127 | 121 | 120 | 121 | 114 | 114 |
| Bank controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Observations | 1,804 | 1,502 | 1,275 | 1,480 | 1,422 | 1,532 | 1,020 | 1,005 |
| $R^2$ | 0.297 | 0.217 | 0.246 | 0.167 | 0.168 | 0.101 | 0.206 | 0.152 |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

All variables are mean-centred and scaled by their respective standard deviation. All independent variables are lagged by 1 quarter, except for columns 7 and 8, where they are lagged by 8 quarters.

### *4.4 Z-score decomposition*

Next, in order to help us understand the channels through which culture affects bank risk, we investigate the relationship between organisational culture and the individual components of the z-score, helping us to understand the channels through which culture affects bank risk. Table 5 provides this breakdown for ROA, leverage, and volatility of ROA for both the summary culture score and the individual cultural dimensions. In Panel A, we see that healthier cultures are associated with better performance, with a coefficient of 0.147, and lower volatility of earnings, with a coefficient of -0.172. When looking at individual dimensions (Panel B of Table 5) we see that only the coefficient on the inclusivity dimension is positive and significant when the dependent variable is ROA. The inclusivity dimension is borderline significant ($p < 0.1$) and negative for the volatility of ROA. Together this suggests that overall organisational culture primarily affects bank risk through the performance and volatility channels. Banks that have healthier cultures tend to also perform better and have more stable earnings.

**Table 5: Regression results for z-score decomposition**

| | Panel A | | |
|---|---|---|---|
| | ROA | Leverage ratio | SD(ROA) |
| | (1) | (2) | (3) |
| Culture | 0.147** | -0.013 | -0.172** |
| | (0.066) | (0.031) | (0.068) |
| Time Fixed effects | Yes | Yes | Yes |
| Clustering level | Banks | Banks | Banks |
| Number of banks | 121 | 121 | 121 |
| Bank controls | Yes | Yes | Yes |
| Observations | 1,020 | 1,019 | 1,018 |
| $R^2$ | 0.135 | 0.531 | 0.226 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

All variables are mean-centred and scaled by their respective standard deviation. All independent variables are lagged 8 quarters.

| | Panel B | | |
|---|---|---|---|
| | ROA | Leverage ratio | SD(ROA) |
| | (1) | (2) | (3) |
| Integrity | 0.028 | -0.033 | -0.004 |
| | (0.031) | (0.024) | (0.019) |
| Detail orientation | 0.121 | -0.024 | -0.102 |
| | (0.081) | (0.038) | (0.091) |
| Customer orientation | -0.075 | -0.023 | -0.066 |
| | (0.050) | (0.026) | (0.048) |
| Inclusivity | 0.217*** | -0.001 | -0.137* |
| | (0.083) | (0.023) | (0.070) |
| Risk orientation | 0.040 | -0.008 | -0.133 |
| | (0.080) | (0.048) | (0.119) |
| Time Fixed effects | Yes | Yes | Yes |
| Clustering level | Banks | Banks | Banks |
| Number of banks | 121 | 121 | 121 |
| Bank controls | Yes | Yes | Yes |
| Observations | 1,020 | 1,019 | 1,018 |
| $R^2$ | 0.177 | 0.536 | 0.236 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

All variables are mean-centred and scaled by their respective standard deviation. All independent variables are lagged 8 quarters.

### *4.5 Endogeneity concerns*

Here, we address the potential endogeneity of our culture score. Although we lagged our explanatory variables, it could be that causality runs in the opposite direction, with bank risk affecting organisational culture. This might be because riskier banks attract individuals with higher risk tolerance, which in turn drives the development of specific organisational cultures. To address this concern, as well as other potential endogeneity issues, we use an instrumental variable (IV) approach and perform coarsened exact matching (CEM) (Corritore et al., 2019; Iacus et al., 2012).

#### *4.5.1 Instrumental variables*

We exploit the presence of the three discrete bank types in our data – whether a bank is a foreign subsidiary of an international bank, a domestic bank, or chartered as a mutual – to instrument for the relationship between culture and risk. A similar instrument was used by Demirgüç-Kunt and Huizinga (2010) when exploring the impact of bank activity and funding strategies on bank risk and return, however in their paper the types were based on business models and in our case the types are rooted in legal and structural differences. For firm type to be a valid instrument, it needs to satisfy the relevance and exogeneity requirements. In other words, firm type should be correlated with organisational culture and not causally related to bank risk.

Regarding relevance, there is evidence suggesting that organisational culture is affected by country of origin – Ashraf and Arshad (2017) finds that a foreign bank's home country culture is more influential on the organisation than the host country. Another study by Berger et al. (2020) shows how national culture is an important determinant of bank failure. Similarly, the mutual form likely influences a firm's culture, typically making them more customer-centric and risk-averse because they do not have short-term pressures from external shareholders (Bholat and Gray, 2013). Below we show the first stage IV regression results with firm type as an explanatory variable and our summary culture score as the dependent variable. Firm type is indeed associated with organisational culture, although there is virtually no difference between foreign and UK-headquartered banks.

Regarding exogeneity, the academic research is divided as to whether customer versus stock or private ownership leads to greater risk. While the rules around funding mix for mutuals (required to be under 50% from wholesale markets) are likely to entail a more secure funding base and lower liquidity risk, a diluted ownership in mutuals (every customer has equal standing) suggests control of managers might be more difficult than non-mutual banks (Fama and Jensen, 1983). Moreover, like banks, mutuals often run into problems for a variety of similar reasons, e.g. uncompetitiveness or poor loan quality, as can be seen by recent failures or near-failures (e.g. Manchester Building Society and The Co-operative Bank) and distressed mergers (e.g. Derbyshire and Cheshire building societies with Nationwide in 2008). Furthermore, decisions around where to form or under what charter have typically been taken decades before our sample starts. While there were a number of demutualisations in the UK in the 1980s and 90s following the 1986 Building Societies Act, none of the firms in our sample undergo a change in type during our period of study.

The second stage IV regression results are in Table 6. The results show that the coefficient for our culture score is significant and far larger in size at 1.27. We fail to reject the hypothesis for the Sargan test of over-identifying restrictions ($p = 0.386$), lending credence to the validity of our instrument. We reject the hypotheses for both the weak instrument and Wu-Hausman tests ($p < 0.01$ in both cases), suggesting the IV regression is more consistent than OLS and that our chosen instrument is not weak.[12]

[12] Regarding test for weak instruments, the reported F statistic for the Stage 1 regression in Table 6 is 53.57, which exceeds the commonly used threshold of 10. However, some recent work by Lee et al. (2020) suggests this threshold might be too small, so we also conduct alternative tests – the Anderson-Rubin (Anderson et al., 1949) and Conditional Likelihood Ratio (Moreira, 2003) tests – which are both rejected ($p < 0.01$), providing additional confidence regarding the strength of the chosen instrument.

**Table 6: Instrumental variable regression results**

| | IV model | |
|---|---|---|
| | Stage 1: Culture | Stage 2: Z-score |
| | (1) | (2) |
| Domestic bank | -0.793*** | |
| | (0.165) | |
| International subsidiary | -0.805*** | |
| | (0.226) | |
| Culture | | 1.270*** |
| | | (0.282) |
| Time Fixed effects | Yes | Yes |
| Clustering level | Banks | Banks |
| Number of banks | 114 | 114 |
| Bank Controls | Yes | Yes |
| Observations | 1,020 | 1,020 |
| Adjusted $R^2$ | 0.309 | 0.225 |
| F Statistic | 53.567*** (df = 9; 992) | 40.273*** (df = 8; 993) |

*Note:* *p<0.1; **p<0.05; ***p<0.01
All continuous variables are mean-centred and divided by their respective standard deviation. Independent variables are lagged 8 quarters.

### *4.5.2 Matching*

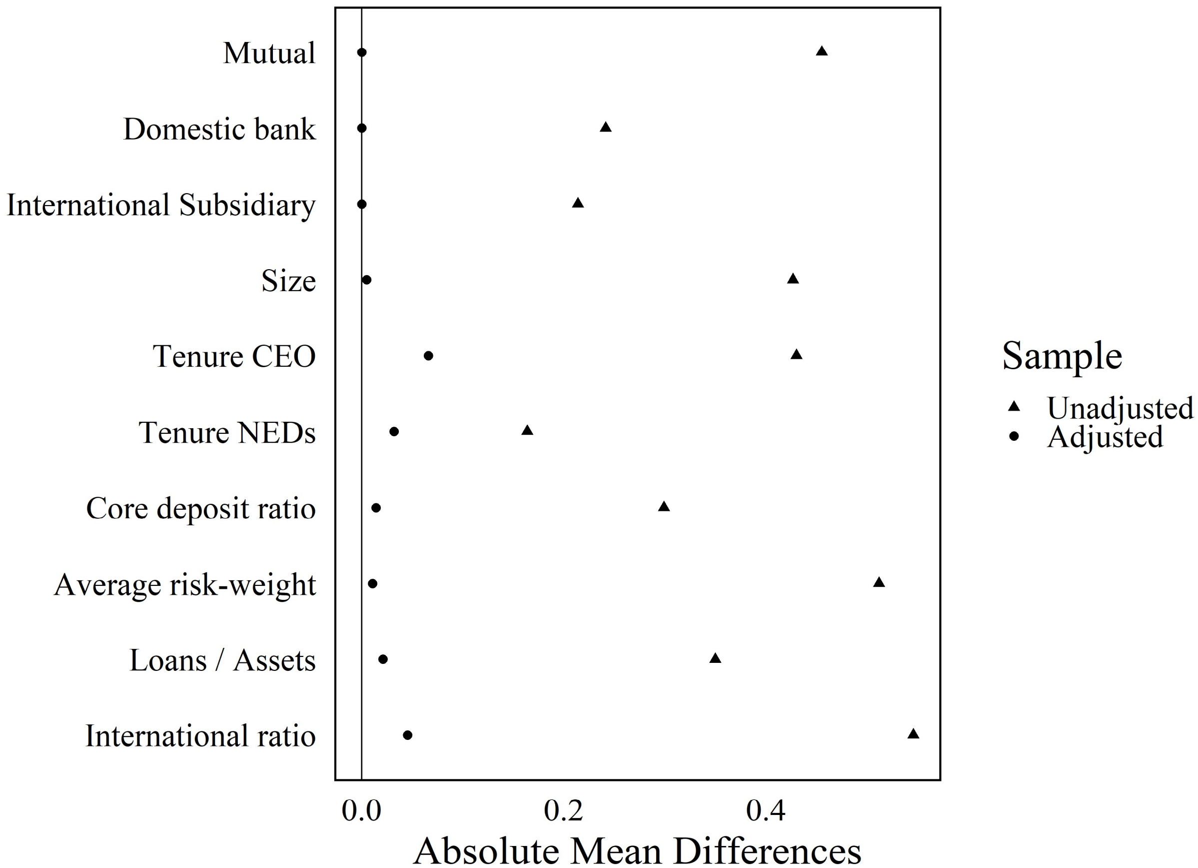
We perform CEM to address potential endogeneity that might arise from selection bias. This approach allows us to match banks exactly on some variables and coarsely on others to achieve balance between treatment and control observations (defined as above or below the global median for culture respectively). We match exactly on the instrument used in the IV estimation, firm type, and coarsely on all the bank-level covariates included in the base analysis. We implemented CEM using the *MatchIt* package in R (Ho et al., 2011). The default binning algorithm did not result in sufficient matched observations, so we manually binned each control variable as follows to achieve balance on the covariates (see Figure 3 for balance assessment pre and post-matching): NED tenure – 6; average risk-weight – 5; size, CEO tenure, and loans to assets – 4; core deposit ratio and international ratio – 3.

**Table 7: Coarsened exact matching regression results**

| | Dependent variable: Z-score | |
|---|---|---|
| | (1) | (2) |
| Culture (above median) | 0.239** | 0.219** |
| | (0.105) | (0.101) |
| Time Fixed effects | Yes | Yes |
| Clustering level | Banks | Banks |
| Number of banks | 90 | 90 |
| Bank controls | No | Yes |
| Observations | 621 | 621 |
| $R^2$ | 0.046 | 0.205 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |
| | All independent variables are lagged 8 quarters. | |

Table 7 shows the results for the CEM analysis. As with our regression models above, we include time fixed effects. We also now include observation weights generated by the matching process, and the standard errors reported in parentheses in Table 7 are clustered both at the bank and subclass (i.e. matched strata) levels. The matching process achieves balance on the covariates according to the Love plot in Figure 3, which demonstrates the absolute mean differences between the treatment and control group pre and post-matching, but at the cost of a reduced sample size. We only have 621 observations for 90 banks remaining. Nevertheless, we find that the coefficient on the culture variable is statistically significant and qualitatively similar when we omit bank-level controls (Column 1) or include them (Column 2), suggesting that being ‘treated’ with above median organisational culture reduces risk by 0.239 and 0.219 of a standard deviation respectively.

**Figure 3: Coarsened exact matching covariate balance assessment**




*Note: Absolute mean differences prior to and after matching adjustments. The balance assessment was carried out with the help of the* `cobalt` *package in R.*

## *4.6 Additional robustness checks*

We carry out a number of additional robustness checks to see whether our results are sensitive to how we construct the summary culture score. First, rather than averaging the UICs, we take the first principal component. Second, we construct our culture score using a simple average but not allowing any missing UICs, i.e. only computing the average for complete cases. For both the principal component and complete case approach, we calculate the measure both with and without the internal fraud UICs due to the relatively smaller number of observations. Third, we construct the culture score by averaging the dimension scores instead of the underlying UICs. We do this first for the case where none of the dimensions values are missing, and then we calculate the culture score by iteratively removing each of the 5 dimensions in turn.

The results for these additional robustness checks are in Table A.2. The coefficient on the culture score for these alternative measures remains significant and in the expected direction throughout.

## 5 Conclusion

Policymakers, financial regulators, and market observers have, for some time, believed organisational culture to be an important factor driving prudential outcomes. The challenge has been how to measure culture effectively. Traditionally, information on bank culture has been gathered from employee self-reports. However, these are known to have limitations, as noted earlier.

What we have done in this paper is marry up a unique conceptual approach with multiple data sources to measure organisational culture unobtrusively. Rather than interrogate bank employees, asking them to 'show us' their bank's culture, we've instead interrogated data that provides insight on different dimensions of culture, from the quality and lateness of their regulatory reporting, to the volume and handling of customer complaints. Importantly, the sources we examine are more directly related to their respective cultural dimension than is typical in prior studies using unobtrusive data. However, we do not have sources of cultural data for every potential dimension of interest. Future work could seek to ascertain such sources of information to paint an even richer picture of culture in UK banking. For example, data on variable remuneration for executives might shed light on the results orientation dimension of culture. In total, we gathered 20 individual indicators of culture pertaining to a large panel of UK banks, ranging from large multi-national firms to small regional building societies.

We find that poor culture leads to greater bank risk as hypothesized. The estimated size of the effect is substantial. These results hold regardless of how we compute the summary culture score or which measure of bank risk we use. Our model is also robust to different subsamples and specifications, and our IV and CEM estimates account for endogeneity concerns, suggesting the relationship between culture and risk is causal.

We think our paper has implications for bank supervisors and policymakers in the UK and beyond. The richness and variety of the data we use can augment more traditional and resource intensive mechanisms for surfacing information on firm cultures via surveys and interviews with senior leaders at firms. Furthermore, the data we have compiled could be used as additional inputs into predictive models of bank risk, potentially rendering these more accurate and practically useful relative to models that rely only financial and macroeconomic data (Suss and Treitel, 2019). Our approach might also offer a blueprint for other sectors and other banking sector stakeholders, for example board committees tasked with monitoring the culture of their institution, or pension and investment funds with an interest in the organisational culture of the banks they invest in.

# Annex

## Table A.1: Definition and source for all indicators of culture

| UIC | Definition | Source |
|---|---|---|
| **Customer orientation** | | |
| Complaints Outstanding | Customer complaints outstanding / opened | FCA Complaints Return |
| Complaints Upheld | Customer complaints upheld / opened | FCA Complaints Return |
| Complaints Closed Slow | Customer complaints closed slowly (longer than 8 weeks) / closed | FCA Complaints Return |
| Redress Paid Per Complaint | Average redress paid / closed complaint | FCA Complaints Return |
| Complaints Opened / Assets | Customer complaints opened / bank assets (£bn) | FCA Complaints Return |
| **Detail orientation** | | |
| Validation Failures | Validation failures (EBA definitions) / regulatory return modules submitted | BOE Regulatory Return Data Quality Metrics |
| Plausibility Flags | Plausibility flags raised (BoE definitions) / regulatory return templates submitted | BOE Regulatory Return Data Quality Metrics |
| Lateness | Late submissions / regulatory return modules submitted | BOE Regulatory Return Data Quality Metrics |
| Late Days | Days submitted late / regulatory returns submitted late | BOE Regulatory Return Data Quality Metrics |
| Early Days | Days submitted early / regulatory returns submitted early | BOE Regulatory Return Data Quality Metrics |
| **Inclusivity** | | |
| Gender Blau | Blau index of female proportion, all authorised positions | Approved Persons Database |
| Gender Difference | Female proportion difference (oversight minus executive) | Approved Persons Database |
| Age S.D. | Age standard deviation, all authorised individuals | Approved Persons Database |
| Age Difference | Age average difference (oversight minus executive positions) | Approved Persons Database |
| **Integrity** | | |
| Operational events: internal fraud | Proportion of operational risk events related to internal fraud | Regulatory returns |
| Operational loss: internal fraud | Proportion of operational risk loss related to internal fraud | Regulatory returns |
| Whistleblowing Referrals | Whistleblowing Referrals | PRA Intelligence |
| **Risk orientation** | | |
| Avg RW Difference | Difference in average risk-weight between bank and division average | Regulatory returns |
| Capital Buffer Difference | Difference in capital buffer between bank and division average | Regulatory returns |
| Pillar 2 / ICG | Pillar 2 capital requirements / total capital requirements | Regulatory returns |

**Table A.2: Alternative culture indices**

| | Alternative summary culture measures | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Principal component (full sample) | Principal component (without internal fraud) | Complete cases (full sample) | Complete cases (without internal fraud) | Average of dimensions | Without integrity | Without detail | Without customer | Without inclusive | Without risk |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Culture | 0.282** | 0.234** | 0.271* | 0.280*** | 0.271*** | 0.217*** | 0.248*** | 0.180*** | 0.225*** | 0.245*** |
| | (0.136) | (0.098) | (0.149) | (0.093) | (0.074) | (0.072) | (0.072) | (0.066) | (0.072) | (0.066) |
| Time Fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Clustering level | Banks | Banks | Banks | Banks | Banks | Banks | Banks | Banks | Banks | Banks |
| Number of banks | 42 | 105 | 42 | 105 | 132 | 136 | 133 | 154 | 132 | 133 |
| Bank controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 96 | 568 | 96 | 568 | 1,020 | 1,263 | 1,072 | 1,493 | 1,024 | 1,047 |
| $R^2$ | 0.360 | 0.235 | 0.366 | 0.256 | 0.199 | 0.176 | 0.200 | 0.155 | 0.180 | 0.192 |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

The estimates are based on alternative measures of organisational culture. The independent variables are lagged 8 quarters.

# References


Anderson, T.W., Rubin, H., others, 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. Annals of Mathematical Statistics 20, 46–63.

Antonsen, S., 2009. Safety culture assessment: A mission impossible? Journal of Contingencies and Crisis Management 17, 242–254.

Ashraf, B.N., Arshad, S., 2017. Foreign bank subsidiaries' risk-taking behavior: Impact of home and host country national culture. Research in International Business and Finance 41, 318–335.

Bailey, A., 2016. Culture in financial services – a regulator's perspective. Speech: 9 May 2016

Berger, A.N., El Ghoul, S., Guedhami, O., Roman, R.A., 2017. Internationalization and bank risk. Management Science 63, 2283–2301.

Berger, A.N., Li, X., Morris, C.S., Roman, R.A., 2020. The effects of cultural values on bank failures around the world. Journal of Financial and Quantitative Analysis 1–49. https://doi.org/10.1017/S0022109020000150

Bholat, D., Gray, J., 2013. Organizational form as a source of systemic risk. Economics 7, 1–35.

Bolino, M.C., Kacmar, K.M., Turnley, W.H., Gilstrap, J.B., 2008. A multi-level review of impression management motives and behaviors. Journal of Management 34, 1080–1109.

Corritore, M., Goldberg, A., Srivastava, S.B., 2019. Duality in diversity: How intrapersonal and interpersonal cultural heterogeneity relate to firm performance. Administrative Science Quarterly 65(2), 359—394.

Demirgüç-Kunt, A., Huizinga, H., 2010. Bank activity and funding strategies: The impact on risk and returns. Journal of Financial Economics 98, 626–650.

Fama, E.F., Jensen, M.C., 1983. Separation of ownership and control. The Journal of Law and Economics 26, 301–325.

Fiordelisi, F., Ricci, O., 2014. Corporate culture and CEO turnover. Journal of Corporate Finance 28, 66–82.

Goldberg, A., Srivastava, S.B., Manian, V.G., Monroe, W., Potts, C., 2016. Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. American Sociological Review 81, 1190–1222.

Gordon, G.G., DiTomaso, N., 1992. Predicting corporate performance from organizational culture. Journal of Management Studies 29, 783–798.

Graham, J.R., Harvey, C.R., Popadak, J., Rajgopal, S., 2017. Corporate culture: Evidence from the field. National Bureau of Economic Research.

Grennan, J., 2019a. Communicating culture consistently: Evidence from banks. Available at SSRN.

Grennan, J., 2019b. A corporate culture channel: How increased shareholder governance reduces firm value. Available at SSRN 2345384.

Guiso, L., Sapienza, P., Zingales, L., 2015. The value of corporate culture. Journal of Financial Economics 117, 60–76.

Guldenmund, F.W., 2007. The use of questionnaires in safety culture research–an evaluation. Safety Science 45, 723–743.

Gupta, A., Owusu, A., 2019. Identifying the risk culture of banks using machine learning. Available at SSRN 3441861.

Heckman, J.J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Hill, A.D., White, M.A., Wallace, J.C., 2014. Unobtrusive measurement of psychological constructs in organizational research. Organizational Psychology Review 4, 148–174.

Ho, D.E., Imai, K., King, G., Stuart, E.A., 2011. MatchIt: Nonparametric preprocessing for parametric causal inference. Journal of Statistical Software 42, 1–28.

Iacus, S.M., King, G., Porro, G., 2012. Causal inference without balance checking: Coarsened exact matching. Political Analysis 1–24.

Leaver, M.P., Reader, T.W., 2019. Safety culture in financial trading: An analysis of trading misconduct investigations. Journal of Business Ethics 154, 461–481.

Lee, D.L., McCrary, J., Moreira, M.J., Porter, J., 2020. Valid t-ratio inference for IV. arXiv preprint arXiv:2010.05058.

Li, K., Mai, F., Shen, R., Yan, X., 2019. Measuring corporate culture using machine learning. Available at SSRN 3256608.

Mahler, J.G., 2009. Organizational learning at NASA: The Challenger and Columbia accidents. Georgetown University Press.

McSweeney, B., 2002. Hofstede's model of national cultural differences and their consequences: A triumph of faith-a failure of analysis. Human Relations 55, 89–118.

Moniz, A., Jong, F. de, 2014. Sentiment analysis and the impact of employee satisfaction on firm earnings, in: European Conference on Information Retrieval. Springer, pp. 519–527.

Moore, C., Detert, J.R., Klebe Treviño, L., Baker, V.L., Mayer, D.M., 2012. Why employees do bad things: Moral disengagement and unethical organizational behavior. Personnel Psychology 65, 1–48.

Moreira, M.J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71, 1027–1048.

Nguyen, D.D., Nguyen, L., Sila, V., 2019. Does corporate culture affect bank risk-taking? Evidence from loan-level data. British Journal of Management 30, 106–133.

Noort, M.C., Reader, T.W., Shorrock, S., Kirwan, B., 2016. The relationship between national culture and safety culture: Implications for international safety culture assessments. Journal of Occupational and Organizational Psychology 89, 515–538.

Nuijts, W., Haan, J. de, 2013. DNB supervision of conduct and culture, in: Financial Supervision in the 21st Century. Springer, pp. 151–164.

O'Reilly III, C.A., Caldwell, D.F., Chatman, J.A., Doerr, B., 2014. The promise and problems of organizational culture: CEO personality, culture, and firm performance. Group & Organization Management 39, 595–625.

O'Reilly III, C.A., Chatman, J., Caldwell, D.F., 1991. People and organizational culture: A profile comparison approach to assessing person-organization fit. Academy of Management Journal 34, 487–516.

Parliament, UK, 2016. Financial services mis-selling: Regulation and redress, House of Commons Committee of Public Accounts.

Power, M., Ashby, S., Palermo, T., 2013. Risk culture in financial organisations: A research report. CARR-Analysis of Risk; Regulation.

Probst, T.M., Graso, M., 2013. Pressure to produce = pressure to reduce accident reporting? Accident Analysis & Prevention 59, 580–587.

Reader, T., Gillespie, A., Hald, J., Patterson, M., 2020. Unobrusive indicators of culture for organizations: A systematic review. European Journal of Work and Organizational Psychology 29, 633–649.

Reader, T.W., Gillespie, A., 2020. Stakeholders in safety: Patient reports on unsafe clinical behaviors distinguish hospital mortality rates. Journal of Applied Psychology.

Reason, J., 2000. Safety paradoxes and safety culture. Injury Control and Safety Promotion 7, 3–14.

Salas, E., Bisbey, T.M., Traylor, A.M., Rosen, M.A., 2020. Can teamwork promote safety in organizations? Annual Review of Organizational Psychology and Organizational Behavior 7, 283–313.

Salz, A., Collins, R., 2013. An independent review of Barclays' business practices.

Sechrest, L., Phillips, M., 1979. Unobtrusive measures: An overview. Unobtrusive measurement today 1–31.

Srivastava, S.B., Goldberg, A., Manian, V.G., Potts, C., 2018. Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations. Management Science 64, 1348–1364.

Suss, Joel, Angeli, Marilena, Eckley, Peter, In Press. Gender and age diversity in UK banks. Bank of England Staff Working Paper.

Suss, Joel, Treitel, Henry, 2019. Predicting bank distress in the UK with machine learning. Bank of England Working Paper No. 831.

Thirty, Group of, 2018. Banking conduct and culture: A permanent mindset change, Group of Thirty Report.

Umphress, E.E., Bingham, J.B., Mitchell, M.S., 2010. Unethical behavior in the name of the company: The moderating effect of organizational identification and positive reciprocity beliefs on unethical pro-organizational behavior. Journal of Applied Psychology 95, 769.

Waring, J.J., 2005. Beyond blame: Cultural barriers to medical incident reporting. Social Science & Medicine 60, 1927–1935.

Webb, E., Campbell, D., Schwartz, R., Sechrest, L., 1966. Unobtrusive measures: Non-reactive research in the social sciences.

Westrum, R., 2004. A typology of organisational cultures. BMJ Quality & Safety 13, ii22–ii27.

Wicker, A.W., 1969. Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. Journal of Social Issues 25, 41–78.

Zhou, X., 2001. Understanding the determinants of managerial ownership and the link between ownership and performance: Comment. Journal of Financial Economics 62, 559–571.