Temperature and the work of bureaucrats

Ana Paula Melo
University of Wisconsin-Madison

Mizuhiro Suzuki*
University of Wisconsin-Madison

June 15, 2021

Abstract

A growing body of literature shows evidence that high temperatures negatively impact performance. Less studied is the impact on the work performance of government bureaucrats. Our paper estimates the impact of temperature on auditors' work performance, measured by the likelihood with which auditors report corruption. We use data on hundreds of municipalities randomly audited in an anti-corruption program in Brazil. We find that auditors are more likely to report corruption if their fieldwork is conducted under higher temperatures. We discuss the potential mechanisms underlying our findings, which highlight important avenues for further research. Our results have implications for understanding the influence of external factors on essential government functions that impact social welfare.

Keywords: temperature, work performance, corruption, government audit
JEL Codes: H83, J81, Q54

*Contact: melodasilva@wisc.edu and msuzuki7@wisc.edu. The authors would like to thank Laura Schechter for her numerous feedback and support. We are also grateful to Raphael Bruce, Joshua Deutschmann, Paul Dower, Jeremy Foltz, Luis Meloni, Andrew Stevens for their helpful comments and suggestions. Any errors in this draft are the sole responsibility of the authors.

1 Introduction

Bureaucratic efficiency is essential in developing countries, where the public sector's effectiveness and efficiency are crucial for successful development activities (World Bank, 2008). Given its importance for the economic growth of low-income countries, what affects public servants' work performance has been an active research area in development economics (Ashraf et al., 2016; Banuri et al., 2018; Dal Bó et al., 2013; Duflo et al., 2012; Glewwe et al., 2010; Kahn et al., 2001; Muralidharan and Sundararaman, 2011). Auditors, in particular, play a critical role in developing economies, where corruption tends to be prevalent.[1] Previous literature provides evidence on how auditors' work impacts electoral results and future corrupt behavior (Avis et al., 2018; Bobonis et al., 2016; Ferraz and Finan, 2008; Olken, 2007; Zamboni and Litschig, 2018).

Despite its importance, we know little about how bureaucrats' performance depends on external work environments. A growing literature suggests seemingly irrelevant factors can affect economic outcomes and distort the consequences of important events. In particular, evidence points towards a loss of efficiency in individual task performance due to high temperature, such as labor productivity (Adhvaryu et al., 2020), work (Lopalo, 2018), and high-stakes exam performance (Graff Zivin et al., 2020; Melo and Suzuki, 2021; Park, 2020). However, the effect of temperature on public servants' work performance is less studied, partly due to the difficulty in measuring their performance.

Our paper fills this gap by analyzing the effect of temperature on auditors' work performance. To measure their performance, we use written audit reports from a federal anti-corruption program in Brazil. Across different months, this program assigned auditors to inspect the use of federal funds at randomly selected municipalities. The economic relevance of temperature for corruption reporting builds on previous findings highlighting the impor-

[1]See Olken and Pande (2012) for a thorough review on this issue.

tance of this anti-corruption program for political outcomes and social welfare. For instance, individuals change their voting behaviors in response to reported municipal corruption levels (Ferraz and Finan, 2008). Audits not only detect corruption ex-post but also deter corrupt behaviors ex-ante (Avis et al., 2018; Bobonis et al., 2016; Olken, 2007; Zamboni and Litschig, 2018).

Our empirical strategy exploits temperature shocks at the time municipalities are audited in the federal audit program. We combine data extracted from the audit reports with high-frequency temperature information during fieldwork to estimate the effects of temperature on auditors' likelihood of reporting corruption. We use corruption indicators compiled and used by Brollo (2011) and Brollo et al. (2013), which classify the reported irregularities as broad and narrow corruption. "Broad" corruption refers to all irregularities, while "narrow" is a subset of the "broad" measure and only includes more severe violations. Additionally, we complement this data with other report elements such as the numbers of figures and tables.

We interpret the estimated relationship of temperature and corruption reporting as the causal effect of temperature on auditors' work performance. The random nature of the audit program guarantees that the municipality assignment and the audit timing are unrelated to weather conditions. Moreover, the federal government's commitment to inspect the selected municipality within a short time does not allow auditors to respond to weather conditions by changing audit schedules. We also control for a range of municipality-level controls, including long-term average temperature, to account for the possibility of a relationship between institutional capacity and the long-run climate.

Results show that temperature has a statistically and economically significant effect on auditors' work performance. We find that a one standard deviation increase in temperature during the audit period increases auditors' likelihood of reporting corruption by 17 percent. Our results are robust to different specifications and alternative performance measurements.

We conduct a falsification test using temperature measurements one year before and one year after the fieldwork. The estimates indicate that temperatures one year before and after the fieldwork have statistically insignificant impacts on corruption reports. This test supports our interpretation that the effects of current temperature on corruption result from temperature shocks during inspections and not a general relationship between long-term climate and corruption. Another concern is that our results are driven by errors in the corruption measures used in our main specification. As an alternative measure of corruption, we collect information on the number of figures and tables in each report. These elements are likely used to support auditors' written corruption claims. We find that the numbers of tables and images also increase if audits are conducted on hotter days. The positive impact of high temperature on the alternative and objective corruption measures supports our finding that temperature affects auditors' work performance.

Municipalities have a chance to revise the audit reports between fieldwork and the report publication, limiting the possibility that higher temperatures induce auditors to misreport more corruption. For instance, municipalities can revise and correct any wrongfully reported corruption before audit results are released to the public. Thus, it is unlikely that corruption misreporting could survive these revision stages. To explain why auditors are more likely to report corruption when the temperature during fieldwork is high, we discuss the following potential mechanisms: (i) the temperature can change how auditors conduct fieldwork and how much information they collect; (ii) the temperature during fieldwork can also affect how they write reports; and (iii) temperature can change how local bureaucrats interact with auditors, affecting, for instance, the type or amount of information available for inspection. Unfortunately, our current dataset does not allow us to test any of these mechanisms. Nonetheless, our discussion opens up topics for future research.

Our study contributes to the literature on the determinants of bureaucrats' work performance. Existing studies have mainly focused on the effect of monetary incentives (Dal Bó et al., 2013; Glewwe et al., 2010; Kahn et al., 2001; Muralidharan and Sundararaman, 2011) and intrinsic motivation (Ashraf et al., 2014; Banuri et al., 2018). Exceptions investigating the temperature effect on public servants' work performance include Heyes and Saberian (2019) and Obradovich et al. (2018), both in the US context. Our main contribution is to provide evidence that temperature can affect bureaucratic efficiency in the high-stakes context of a developing country. Our findings also contribute to the emerging literature on the short-term effect of external factors on economic outcomes, in contrast to the studies on the long-run impacts of climate on economic outcomes (Burke et al., 2015; Carleton and Hsiang, 2016; Dell et al., 2012). Examples of such external factors include temperature (Heal and Park, 2016), pollution (Ebenstein et al., 2016), and noise (Dean, 2019).

This study proceeds as follows. In Section 2, the institutional background of the anti-corruption program in Brazil is introduced. Section 3 describes the data. Our empirical strategy is introduced in Section 4. In Section 5, we present the empirical results and discuss the potential mechanisms underlying the results. Section 6 concludes.

2 Background: the anti-corruption program

In 2003, the Brazilian federal government started a national anti-corruption program to investigate local corruption at the municipality level. In this program, municipalities are randomly selected by a public lottery, and municipal governments' expenditures of federal transfers are audited. In each round, 50 or 60 municipalities were randomly drawn, with replacement, from all municipalities with fewer than 450,000 residents. After a municipality is chosen in a lottery, the Controladoria Geral da União (CGU), a branch of the federal government in charge of transparency policies, sends around 10 to 15 auditors to each selected municipality.

Each state typically has its local *controladoria* branch, which oversees the audits in its municipalities. Auditors investigate evidence of corruption or general misconduct in the use of federal resources. They also interview local community and municipal council members. Audits typically take five days, from Monday to Friday, and afterward, auditors write a report about irregularities and evidence of corruption found during the fieldwork. The municipality administration can respond to irregularities found, and the audit team registers in the report whether the justifications were accepted or not. Reports are sent to relevant areas of the federal government, often responsible for the federal resources in which irregularities were found. This report is also made public on the CGU website, which is the information we analyze. Summaries of these reports are communicated through media outlets, which voters pay attention to and can, for instance, change their voting behaviors (Ferraz and Finan, 2008).

It is important to highlight that the auditors' career is competitive with high salaries, lowering incentives to accept bribes.[2] Additionally, they receive extensive training before being sent to inspect municipalities. Auditors receive guidelines on how to produce reports, with standardized rules about formatting and content. Reports often contain photo images taken during the fieldwork or tables made based on the information collected during the fieldwork.

Overall, the nature of their work, which is defined by intensive workdays and field visits, can be affected by local weather conditions during the fieldwork. They can affect auditors' ability to interview local community members or oversee development projects, potentially impacting their ability to report more or less corruption. The potential effects of these external conditions on the probability of reporting corruption are the main object of interest in our paper.

[2]According to interviews by Ferraz and Finan (2008), program officials had never seen an incident of offered side payments to auditors. The authors argue that this could be because mutual monitoring within audit teams works as a deterrent to bribes.

3 Data

We use three sets of data for the analyses: weather, reported corruption measures, and municipality characteristics. First, for weather-related measures, we use the Princeton Global Meteorological Forcing Dataset. This is a gridded (0.25-degree latitude/longitude), three-hourly dataset, which allows us to access weather information in developing countries where weather stations exist only sparsely. Its temporal resolution makes it possible for us to use average temperature between 9 am to 5 pm, roughly corresponding to the time auditors work in the field.

This dataset contains information on rainfall, air pressure, specific humidity, and dry-bulb temperature, which is the temperature that one would often refer to in daily life. With these variables, we calculate wet-bulb temperature, which is the temperature measure we use in our main results. As opposed to dry-bulb temperature, wet-bulb temperature takes into account the air temperature and humidity simultaneously. This measure has been extensively used in climate science and biology to represent heat stress danger and thermal comfort (Budd, 2008; Liljegren et al., 2008). It has been used in recent economics studies as well (Adhvaryu et al., 2020; Geruso and Spears, 2018). A higher humidity prevents the body from cooling itself through evaporative heat loss from the skin. Therefore, wet-bulb temperature affects human comfort, potentially impacting how auditors conduct fieldwork and write audit reports.[3] The distributions of temperature measures are shown in Figure C.1.

Secondly, as a measure of reported corruption, we use data manually compiled by Brollo (2011) and Brollo et al. (2013) and made available online for other researchers. The authors coded information contained in the publicly available audit reports. Broad corruption in-

[3]For more details about this data and how to calculate wet-bulb temperature, refer to Appendix A. The results using dry-bulb temperature and the differences from the estimates using wet-bulb temperature are discussed in the Results section.

cludes "irregularities that could also be interpreted as bad administration rather than overt corruption." Therefore, all reported irregularities are classified as broad corruption.[4] A subset of these irregularities is classified as narrow corruption, defined as "only severe irregularities that are also more likely to be visible to voters." For more details about the definition of corruption measures, see Brollo et al. (2013). As outcome variables, we use the indicators of whether auditors reported any broad and narrow corruption for audited municipalities.

We complement this dataset by collecting additional information from the audit reports. The most important information is the fieldwork dates. Although this corruption dataset contains information from municipalities audited in the 2nd to 29th rounds, we restrict our analyses to the 2nd-19th audit rounds to increase the precision of our targeted fieldwork dates. For instance, between the 2nd to 19th audit rounds, more than 99 percent of the reported fieldwork length is less than or equal to 12 days. After the 19th round, the duration of the fieldwork informed in the reports is much longer, often lasting over one month. When the reported fieldwork is too long, there is considerable uncertainty about when the actual fieldwork happened. We also collect information on the total number of figures and tables in each report as additional information on corruption and public resource mismanagement.

Thirdly, we obtain municipality characteristics from the Population Census in 2000, which we add as control variables in our regressions. These include the urban population share, income per capita, log of population, and population share below the poverty line.

The final data contains 966 audits conducted from the 2nd to 19th audit rounds, from 2003 to 2005.[5] Summary statistics are shown in Table 1, by type of irregularity. Auditors reported broad corruption in 711 audits, of which 424 were also classified as narrow corrup-

[4]According to this definition, this measure also includes mismanagement of public resources. Following the terminology in Brollo et al. (2013), we call it broad corruption.

[5]Since municipalities are randomly selected with replacement, 11 municipalities are audited twice during our sample period. In our analysis, we control for an indicator for whether a municipality is audited for the second time.

tion. The table indicates that audits resulting in reported corruption are more likely to be conducted on hotter days. Additionally, more images and tables are used in audit reports in which corruption is reported. This suggests that auditors use images and tables as evidence of corruption and mismanagement.

Table 1: Summary Statistics (by corruption definitions)

	(1)	(2)	(3)	(4)	(5)	(6)
	Broad = 1	Broad = 0	Broad Diff. ((1) - (2))	Narrow = 1	Narrow = 0	Narrow Diff. ((4) - (5))
Wet bulb temp. (°C)	23.30	22.15	1.15***	23.57	22.56	1.01***
	(3.31)	(4.04)	(0.26)	(3.19)	(3.75)	(0.23)
Dry bulb temp. (°C)	27.97	26.44	1.53***	28.18	27.08	1.10***
	(3.96)	(4.89)	(0.31)	(3.80)	(4.56)	(0.28)
Number of images	19.81	18.14	1.67	22.60	16.84	5.76***
	(22.38)	(24.56)	(1.68)	(24.08)	(21.75)	(1.48)
Number of tables	27.84	19.08	8.76***	32.66	19.95	12.71***
	(23.42)	(13.62)	(1.55)	(26.50)	(14.60)	(1.34)
Rainfall (mm/day)	2.74	3.26	-0.52*	2.70	3.02	-0.32
	(3.48)	(4.59)	(0.28)	(3.46)	(4.06)	(0.25)
Long-run average wet bulb temp.	24.26	23.56	0.71***	24.45	23.79	0.66***
	(2.32)	(2.71)	(0.18)	(2.24)	(2.56)	(0.16)
Long-run average dry bulb temp.	28.30	27.28	1.01***	28.49	27.67	0.81***
	(2.84)	(3.31)	(0.22)	(2.72)	(3.16)	(0.19)
Long-run average rainfall	3.81	4.33	-0.52***	3.72	4.13	-0.41***
	(1.39)	(1.29)	(0.10)	(1.47)	(1.29)	(0.09)
Share of pop urban (%)	59.23	59.55	-0.32	57.78	60.52	-2.74*
	(22.10)	(24.86)	(1.67)	(21.94)	(23.49)	(1.48)
Log pop.	9.50	9.47	0.04	9.55	9.45	0.10
	(0.98)	(1.12)	(0.07)	(0.99)	(1.04)	(0.07)
Share of pop poor (%)	44.80	37.41	7.39***	48.51	38.43	10.09***
	(22.50)	(22.00)	(1.63)	(22.31)	(21.84)	(1.43)
Second-time audit	0.01	0.01	0.00	0.01	0.01	-0.01
	(0.11)	(0.09)	(0.01)	(0.08)	(0.12)	(0.01)
Observations	711	255	966	424	542	966

Notes: In columns (1), (2), (4), and (5), means (and standard deviations) are shown. In columns (3) and (6), differences in variable means between municipalities with and without corruption detected (and their standard errors) are shown. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

4 Empirical Strategy

Our objective is to identify the effect of temperature during fieldwork on the probability with which auditors report corruption. Descriptive statistics reveal a positive correlation between temperature and reported corruption. To identify the causal effect of temperature on the probability of reporting corruption, we exploit the exogenous variation in temperature when audits are conducted. Regression equations are displayed below, in equations (1) and (2).

The outcome variable, C_{ijt}, is an indicator for any corruption (broad and narrow separately) reported for a municipality i in a mesoregion j that is audited in the t'th round of the audit program.[6] To control for confounding factors, we use average rainfall during fieldwork (mm/day) (P_{ijt}) and municipality-level variables (X_{ij}).[7] We also use mesoregion fixed effects (μ_j) and audit wave fixed effects (ν_t). The error term is denoted by ϵ_{ijt}.

We use two regression specifications. In the first specification, the effect of wet-bulb temperature (T_{ijt}) is assumed to affect the outcome linearly. The second specification allows a non-linear relationship between temperature and reported corruption. In this more flexible specification, we use the binned wet-bulb temperature (T^b_{ijt}) with 2°C-intervals.[8]

$$C_{ijt} = \alpha T_{ijt} + \beta P_{ijt} + X'_{ij}\gamma + \mu_j + \nu_t + \epsilon_{ijt} \tag{1}$$

$$C_{ijt} = \sum_b \alpha_b T^b_{ijt} + \beta P_{ijt} + X'_{ij}\gamma + \mu_j + \nu_t + \epsilon_{ijt}. \tag{2}$$

Equation (1) assumes a linear effect of temperature, while Equation (2) allows a more flexible relationship between temperature and the outcome. To account for correlations within states due to, for example, audit-team specific factors, inferences are based on the clustered wild-bootstrap method at the state level (Cameron et al., 2008).[9]

Four points are worth noting. First, the positive relationship between temperature and corruption can be due to other factors such as long-term climate affecting institutions in

[6]Mesoregion is an administrative unit smaller than state, and there are 137 mesoregions in Brazil.

[7]The following covariates are included in the regressions: municipality long-run temperature, the share of the urban population, income per capita, the log of population, population share below the poverty line, and an indicator for the second-time audit.

[8]We select 2°C-intervals so that all temperature bins contain at least around 10 percent of observations while maximizing the number of equal-sized bins.

[9]There are 26 states and one federal district in Brazil.

municipalities, as argued by Dell et al. (2012). The inclusion of municipality-level long-run temperature in X_{ij} aims to deal with such concern. Secondly, mesoregion fixed effects account for mesoregion-level climate characteristics such as variance in the long-run temperature. They rule out the possibility that a potentially complex mesoregion-level relationship between climate and institution biases our estimates. Mesoregion fixed effects also account for differences across states, such as differences across state-based audit teams.[10] Thirdly, since only 11 municipalities receive multiple audits during our study period, we cannot use municipality fixed effects. The possibility that mesoregion fixed effects do not control for municipality-level climate characteristics is discussed in the falsification test below. Finally, since audits are conducted just after monthly audit lotteries, we add audit wave fixed effects to control for seasonality.

5 Results

5.1 Main results

The results from the linear specification are shown in Table 2, columns (1) and (2). We show 90 percent confidence intervals calculated based on the clustered wild bootstrap method in square brackets.[11] The table shows the positive effects of wet-bulb temperature on the probability of reporting corruption. The point estimates are not only statistically significant but also economically meaningful: a one standard deviation increase in wet-bulb temperature increases the probability of reporting narrow corruption by 7.5 percentage points (or 17 percent). Since rainfall has only negligible and statistically insignificant effects on the

[10] As a robustness check, we use state fixed effects instead of mesoregion fixed effects. The results are robust to this change and are shown in the appendix.

[11] Standard errors of the estimates are not shown since using standard errors derived from the standard deviations of the bootstrap distribution for inference depends on the asymptotic normality of the estimates (Cameron et al., 2008; Roodman et al., 2019). With the state-level clustered standard errors not based on the wild bootstrap method, Table B.1 provides quantitatively similar and statistically significant results.

outcomes, we focus on temperature effects in the discussions below.

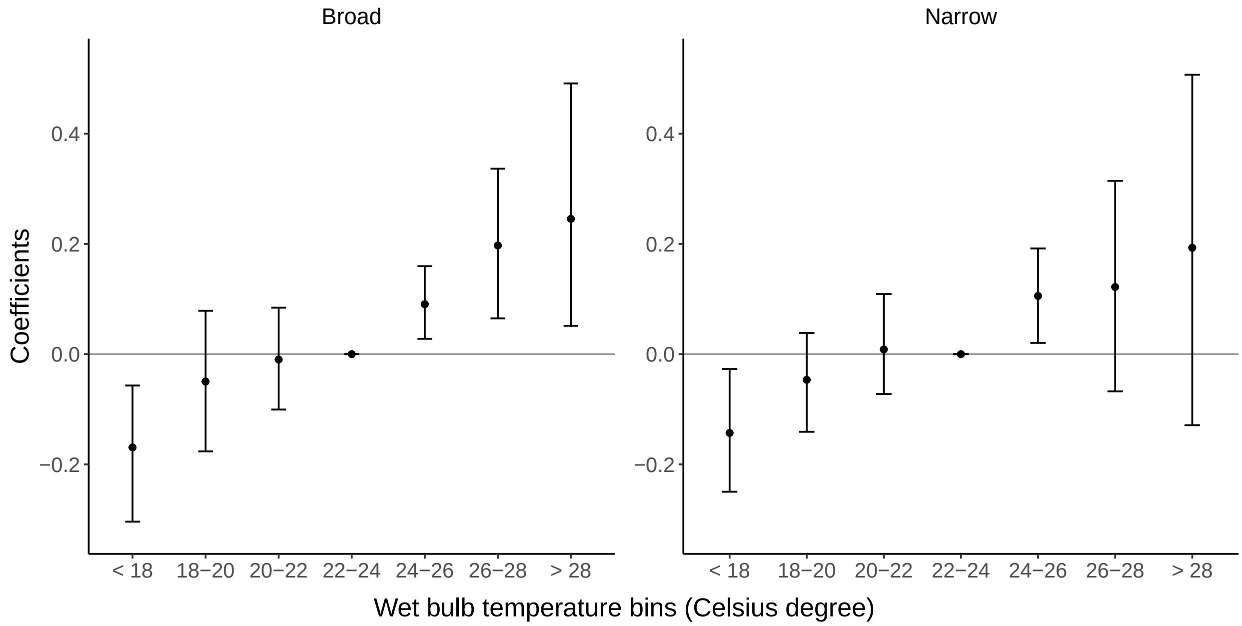
Figure 1 shows the regression coefficients allowing non-linear effects of wet-bulb temperature. Consistent with the linear specification results, the figure shows that auditors are more likely to report corruption when the temperature is higher.

Table 2: Regression: Corruption reports and temperature

	(1) Broad	(2) Narrow	(3) Broad	(4) Narrow
Wet bulb temp. (°C)	0.017	0.021	0.021	0.023
	[0.001, 0.031]*	[0.002, 0.037]*	[0.002, 0.039]*	[0.003, 0.040]*
Wet bulb temp. (one-year lag) (°C)			-0.011	-0.004
			[-0.028, 0.008]	[-0.021, 0.016]
Wet bulb temp. (one-year lead) (°C)			-0.001	-0.001
			[-0.023, 0.016]	[-0.029, 0.023]
Rainfall (mm/day)	0.000	0.001	0.001	0.001
	[-0.010, 0.011]	[-0.007, 0.011]	[-0.010, 0.013]	[-0.006, 0.011]
Observations	966	966	966	966
R-squared	0.27	0.32	0.27	0.32
Mesoregion FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean of outcome	0.736	0.439	0.736	0.439
SD of temperature in year t	3.550	3.550	3.550	3.550

Notes: The dependent variable in columns (1) and (3) is an indicator for broad corruption reported, and the dependent variable in columns (2) and (4) is an indicator for narrow corruption reported. Control variables include the share of the population that is urban, income per capita, log of population, population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Numbers in square brackets are 90% confidence intervals, calculated with clustered wild-bootstrap at the state level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Figure 1: Wet-bulb temperature effects on reported corruption



Notes: The figure shows the regression coefficients of an indicator for corruption (left: broad, right: narrow) reported on the wet-bulb temperature bins. Each temperature interval does not include the right end. The 90% confidence intervals are shown. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Audit wave fixed effects and mesoregion fixed effects are included in the regressions. Confidence intervals are calculated with clustered wild-bootstrap at the state level.

The difference between the results of the wet-bulb and dry-bulb temperatures is worth discussing. Table B.2 in the appendix shows positive but statistically insignificant effects of dry-bulb temperature on corruption reports. Furthermore, Figure C.2 shows the *negative* impact of high dry-bulb temperature (higher than 32°C) on reporting broad corruption. The difference from regression results with wet-bulb temperature suggests the importance of taking humidity into account in this context. In particular, the negative impact of high dry-bulb temperature found in Figure C.2 might be capturing the effect of lower humidity. Indeed, due to the low humidity level, the average wet-bulb temperature for municipalities in the highest dry-bulb temperature bin (> 32) is 25.4°C, which is lower than that of the second-highest dry-bulb temperature bin ($30 - 32$), which is 26.1°C. This resonates with the role of humidity affecting individual comfort found in previous studies (Jing et al., 2013; Li et al., 2019). Moreover, the results are consistent with findings in economics studies such as Geruso and Spears (2018) and Lopalo (2018), which find an important interactive effect between temperature and humidity on health and labor outcomes.

5.2 Robustness checks

5.2.1 Falsification test

In the analyses above, we argue that the relationship between temperature and corruption reports is causal. One concern in this argument is that municipalities' long-run climate characteristics cause spurious correlations between temperature during fieldwork and corruption reports. Although we control for the long-run temperature and rainfall of municipalities as well as mesoregion fixed effects, these may not be sufficient to capture the potentially complex relationship between climate and corruption in each municipality.

To deal with the possibility for spurious correlation, as a falsification test, we add, to Equation (1), temperature measures on the same days one year before and one year after the actual fieldwork. This allows us to test whether results in columns (1) and (2) in Table 2 are spurious correlations caused by the relationship between the long-run weather and actual corruption levels. If the relationship is not spurious, we would observe a statistically insignificant relationship between corruption reports and the lead and lag variables.

The results from the falsification test are shown in columns (3) and (4) in Table 2. They show negligible effects of past and future temperature and unchanged point estimates of temperature measures during the actual fieldwork. These reinforce our argument that we identify the contemporaneous temperature effects on the probability of reporting corruption.[12]

[12]For the reported falsification test, we estimate a modified version of Equation (1) that includes contemporaneous, lag, and lead temperatures. We test whether the coefficients associated with the lag and lead temperature measures are statistically significant. This is our preferred specification since it allows us to examine how the estimate on current temperature changes by including presumably unimportant lag and lead

5.2.2 Regression results with numbers of images and tables in reports

To further corroborate our findings, we use the contents of audit reports and analyze how they are affected by temperature during fieldwork. Specifically, we obtain the numbers of images and tables used in audit reports and use them as outcome variables in regressions. Since, in many cases, auditors use images and tables as proof of corruption or mismanagement, we consider them as a proxy of reported corruption. Positive correlations between reported corruption and the number of figures and tables support this view (Table 2). Appendix D shows several examples of images and tables in audit reports and discusses how they are used as evidence of corruption and bad administration.

The objectivity of figures and tables as a measure of corruption is an advantage over the corruption measures used in our main analysis. Since the "narrow" and "broad" corruption measures are manually compiled, they can contain measurement errors due to subjective judgment during data entry and classification. Note that our proxies can also contain a different type of measurement error. Not all images and tables collected during fieldwork are necessarily used in the reports as proofs of corruption. We rely on the consistent results across different reported corruption measures to support the robustness of our findings.

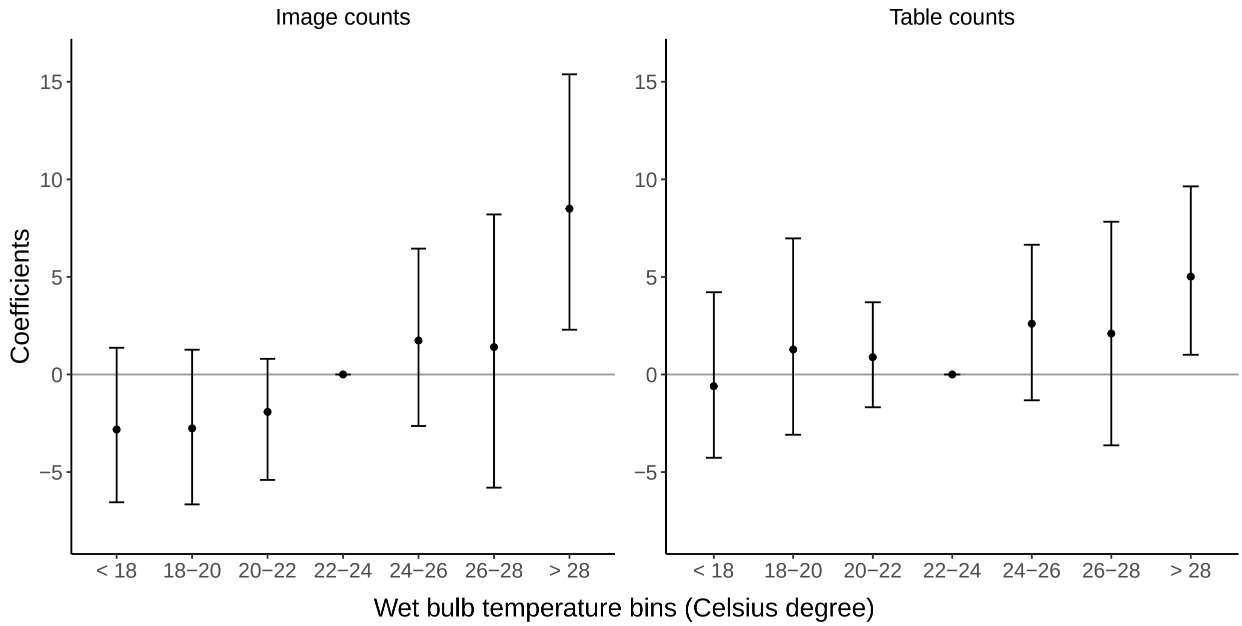
Consistent with our main results, Figure 2 shows the positive effect of high wet-bulb temperature on the numbers of images (left panel) and tables (right panel).[13] These results suggest that when the temperature during fieldwork is higher, auditors report more corruption using visual evidence (e.g., figures and tables). It should be noted that the results using figures and tables do not necessarily reflect an effect of temperature on how auditors conduct fieldwork. While images and tables are more likely to be included in reports if there

temperatures. As an alternative falsification test, we also test whether only lag and lead temperatures, added one at a time, affect reported corruption. Table B.3 in the appendix shows statistically and economically insignificant effects of these temperatures, which is aligned to our preferred test.

[13]Results with a linear specification in Table B.4 show imprecise and statistically insignificant estimates. This can be because a linear specification did not fully capture the non-linear nature of the relationship between temperature and reported figures/tables. Figure C.3 shows overall positive but weaker effects of dry-bulb temperature, which is consistent with the discussion in Section 5.1.

is corruption or mismanagement involved, auditors may not use collected evidence if no issue is found.

Figure 2: Wet-bulb temperature effects on number of images and tables



Notes: The figure shows the regression coefficients of the number of images (left) and the number of tables (right) on the wet-bulb temperature bins. Each temperature interval does not include the right end. The 90% confidence intervals are shown. Control variables include the urban population share, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Audit wave fixed effects and mesoregion fixed effects are included in the regressions. Confidence intervals are calculated with clustered wild-bootstrap at the state level.

5.2.3 Results with different geographic fixed effects

In our primary analyses, we use mesoregions as geographic fixed effects. Since mesoregion is an administrative unit smaller than state, mesoregion fixed effects allow us to control for state-specific factors, such as the audit teams' effect. It also controls for other factors potentially confounding our results, such as long-run climate in each mesoregion. To check the robustness of our decision over the level of geographic fixed effects, we run regressions with state fixed effects instead. Results using state-fixed effects are qualitatively similar to those using mesoregion fixed effects (Table B.5 and Figures C.4 and C.5).

5.3 Discussion

Our work provides evidence that external factors affect bureaucrats' work performance. Specifically, our results indicate that auditors are more likely to report corruption if the temperature is higher during audit fieldwork. It is important to note that municipalities have the chance to revise, explain, and correct any irregularities wrongfully flagged before the public release of the reports. Hence, any corruption that is reported without solid evidence may not survive the municipalities' revision stage. Therefore, the positive effects of temperature on corruption reporting are unlikely to be due to corruption over-reporting.

Due to data limitations, we cannot test for any specific mechanisms behind the relationship between temperature and corruption reporting. Instead, we discuss the following potential drivers of our findings: (i) temperature affects how auditors conduct fieldwork; (ii) temperature during fieldwork affects how auditors write reports; and (iii) temperature affects how local bureaucrats or community members interact with auditors.

First, the temperature can change how auditors conduct fieldwork. We explore two channels through which these effects might operate: a change in mood and a change in productivity. The mood channel could explain why higher temperature causes auditors to detect and report more corruption. Whereas there have been studies on the effect of weather on violent behaviors (Card and Dahl, 2011; Kenrick and MacFarlane, 1986; Larrick et al., 2011), recent studies have also identified the causal impact of temperature on non-violent outcomes related to emotion and mood. For instance, with sentiment extracted from text data on Twitter, Baylis (2020) finds that extreme temperature worsens the expressed sentiment and increases the frequency of aggressively profane phrases in tweets. Similar effects are confirmed in Baylis et al. (2018) with text data on Facebook posts. Along the same line, Heyes and Saberian (2019) find that when the temperature on a case day is higher, judges are more likely to make decisions unfavorable to applicants. They attribute a part of the mechanism to mood: temperature affects irritation and comfort, affecting court judges' decision-making.

On the other hand, a productivity channel could explain our results if auditors under-report corruption when the temperature is lower. Temperature effects on productivity have been estimated in other contexts as well. Their findings suggest both high and low temperatures can affect work performance. For instance, whereas Adhvaryu et al. (2020) find lower productivity under higher temperature, Stevens (2017) finds that lower temperature decreases labor productivity. Although we cannot test this hypothesis, this could explain our findings if *lower* temperature decreases auditors' productivity and, as a result, they collect less evidence of corruption.

Secondly, the temperature during fieldwork can change how auditors write reports, which, in turn, can change the probability of reporting corruption. If auditors write reports at the audited municipalities during or after their fieldwork, the temperature while writing the reports can affect the mood (Baylis, 2020; Baylis et al., 2018; Heyes and Saberian, 2019) and cognitive ability (Graff Zivin et al., 2020; Park, 2020) of auditors, which can change the contents of the reports. Additionally, the effects of temperature *during fieldwork* on report writing could last beyond fieldwork. Previous works suggest that if high temperature induces stress during their fieldwork, this could affect what auditors remember when writing reports. For instance, Hoscheidt et al. (2014) experimentally find that negative memories encoded under stress retain for more extended periods. Therefore, even if auditors write reports elsewhere, such as from their state office, the temperature during fieldwork can still affect report writing. We cannot precisely test if these mechanisms work in our context since we do not have information on where reports are written.

Thirdly, the temperature may affect the interaction between auditors and local bureaucrats. As described before, to obtain information on corruption and mismanagement, auditors interview local community and municipal council members. The effects of temperature through mood or productivity could also impact local bureaucrats, potentially altering the information or evidence provided to auditors.

Investigating the exact mechanism behind our findings is beyond the scope of this study. With our currently available data, we cannot confirm or rule out any of these potential mechanisms. For this, we would need more detailed data on auditors and how they work and write reports. For instance, to analyze the effect of temperature on how auditors work in the field, it is required to collect detailed information on daily activities and complete information on what evidence they collect.[14] Furthermore, for the effect of temperature on how reports are written, it is necessary to obtain information on when and where auditors write reports, how the reports are revised, and how municipalities respond before reports go public. Finally, more information on local bureaucrats' roles during the fieldwork is necessary to provide evidence on how temperature can affect their interaction with auditors. Since none of the information is available to us, such analyses are left for future research.

Finally, our results and discussion suggest that mitigating the temperature effect could enhance the consistency and effectiveness of the audits. The specific program we study has shown positive and lasting effects in reducing corruption in Brazil. As Avis et al. (2018) find, audits in this context work as a deterrent to corruption. Audits increase the perceived costs of corruption and, as a result, decrease corruption. Similar effects of audits have been found in other contexts as well (Bobonis et al., 2016; Olken, 2007; Zamboni and Litschig, 2018). However, this important deterrent effect can be diminished if audit results are prone, for example, to misreporting due to external factors. Understanding the mechanisms behind our results is crucial to provide public policy alternatives.

[14]Notice that this is different from the information auditors decide to include in their reports, which we partially use in the analysis with the numbers of images and tables.

6 Conclusion

In this study, we investigate the effects of seemingly irrelevant factors on the productivity of bureaucrats. Using random municipality assignment to audits and temperature shocks, we estimate the effect of temperature on corruption reports. We find that higher temperature during fieldwork increases the probability of reporting corruption. An increase by one standard deviation in wet-bulb temperature increases the probability of reporting corruption by 17 percent. Our falsification test using lag and lead temperatures reinforces our argument that this relationship is causal and not confounded by such factors as the relationship between long-term climate and institutions. Furthermore, we use the numbers of images and tables in the audit reports and find positive temperature effects on these measures, supporting our results' robustness.

Our findings suggest that work conditions might influence high-stakes audits conducted at specific times. Improved work conditions or specific training to raise awareness of these potential external factors can mitigate the identified effects. However, understanding the mechanisms through which temperature affects auditors' work performance can provide accurate guidance on the most effective type of intervention. We cannot fully tackle this point in our work, and investigating the mechanisms is left for future research.

It is also worth exploring alternative ways to identify corruption that do not rely solely on measures sensitive to the influence of short-run external factors. For instance, repeated audits at different months can smooth out temperature shocks' effects on a specific audit date. Given the potential impact of audits on electoral outcomes and social welfare, taking measures to reduce misreporting is desirable.

References

Adhvaryu, A., Kala, N., and Nyshadham, A. (2020). The light and the heat: Productivity co-benefits of energy-saving technology. *The Review of Economics and Statistics*, 102(4):779–792.

Ashraf, N., Bandiera, O., and Jack, B. K. (2014). No margin, no mission? A field experiment on incentives for public service delivery. *Journal of Public Economics*, 120:1–17.

Ashraf, N., Bandiera, O., and Lee, S. S. (2016). Do-gooders and go-getters: Selection and performance in public service delivery. mimeo.

Avis, E., Ferraz, C., and Finan, F. (2018). Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians. *Journal of Political Economy*, 126(5):1912–1964.

Banuri, S., Keefer, P., and de Walque, D. (2018). *Love the job... or the patient? Task vs. mission-Based motivations in health care*. The World Bank.

Baylis, P. (2020). Temperature and temperament: Evidence from Twitter. *Journal of Public Economics*, 184:104161.

Baylis, P., Obradovich, N., Kryvasheyeu, Y., Chen, H., Coviello, L., Moro, E., Cebrian, M., and Fowler, J. H. (2018). Weather impacts expressed sentiment. *PloS ONE*, 13(4):e0195750.

Bobonis, G. J., Cámara Fuertes, L. R., and Schwabe, R. (2016). Monitoring corruptible politicians. *American Economic Review*, 106(8):2371–2405.

Brollo, F. (2011). Why do voters punish corrupt politicians? Evidence from the Brazilian anti-corruption program. mimeo.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The political resource curse. *American Economic Review*, 103(5):1759–1796.

Budd, G. M. (2008). Wet-bulb globe temperature (WBGT) - Its history and its limitations. *Journal of Science and Medicine in Sport*, 11(1):20–32.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235–239.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Carleton, T. A. and Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, 353(6304).

Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics*, 128(3):1169–1218.

Dean, J. T. (2019). Noise, cognitive function, and worker productivity. mimeo.

Dell, M., Jones, B. F., and Olken, B. a. (2012). Climate shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3):66–95.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–78.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8(4):36–65.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, 123(2):703–745.

Geruso, M. and Spears, D. (2018). Heat, humidity, and infant mortality in the developing world. Technical report, National Bureau of Economic Research, Cambridge, MA.

Glewwe, P., Ilias, N., and Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics*, 2(3):205–27.

Graff Zivin, J., Song, Y., Tang, Q., and Zhang, P. (2020). Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China. *Journal of Environmental Economics and Management*, 104:102365.

Heal, G. and Park, J. (2016). Reflections - Temperature stress and the direct impact of climate change: A review of an emerging literature. *Review of Environmental Economics and Policy*, 10(2):347–362.

Heyes, A. and Saberian, S. (2019). Temperature and decisions: Evidence from 207,000 court cases. *American Economic Journal: Applied Economics*, 11(2):238–265.

Hoscheidt, S. M., LaBar, K. S., Ryan, L., Jacobs, W. J., and Nadel, L. (2014). Encoding negative events under stress: High subjective arousal is related to accurate emotional memory despite misinformation exposure. *Neurobiology of Learning and Memory*, 112:237–247.

Jing, S., Li, B., Tan, M., and Liu, H. (2013). Impact of relative humidity on thermal comfort in a warm environment. *Indoor and Built Environment*, 22(4):598–607.

Kahn, C. M., Silva, E. C., and Ziliak, J. P. (2001). Performance-based wages in tax collection: The Brazilian tax collection reform and its effects. *The Economic Journal*, 111(468):188–205.

Kenrick, D. T. and MacFarlane, S. W. (1986). Ambient temperature and horn honking: A field study of the heat/aggression relationship. *Environment and Behavior*, 18(2):179–191.

Larrick, R. P., Timmerman, T. A., Carton, A. M., and Abrevaya, J. (2011). Temper, temperature, and temptation: Heat-related retaliation in baseball. *Psychological Science*, 22(4):423–428.

Li, C., Liu, H., Li, B., and Sheng, A. (2019). Seasonal effect of humidity on human comfort in a hot summer/cold winter zone in China. *Indoor and Built Environment*, 28(2):264–277.

Liljegren, J. C., Carhart, R. A., Lawday, P., Tschopp, S., and Sharp, R. (2008). Modeling the wet bulb globe temperature using standard meteorological measurements. *Journal of Occupational and Environmental Hygiene*, 5(10):645–655.

Lopalo, M. (2018). Temperature, worker productivity, and adaptation : Evidence from survey data production. mimeo.

Melo, A. P. and Suzuki, M. (2021). Temperature, effort, and achievement: Evidence from a large-scale standardized exam in Brazil. mimeo.

Muralidharan, K. and Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy*, 119(1):39–77.

Obradovich, N., Tingley, D., and Rahwan, I. (2018). Effects of environmental stressors on daily governance. *Proceedings of the National Academy of Sciences*, 115(35):8710–8715.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115(2):200–249.

Olken, B. A. and Pande, R. (2012). Corruption in developing countries. *Annual Review of Economics*, 4(1):479–509.

Park, R. J. (2020). Hot temperature and high stakes performance. *Journal of Human Resources*, forthcoming.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., and Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal: Promoting communications on statistics and Stata*, 19(1):4–60.

Sheffield, J., Goteti, G., Wood, E. F., Sheffield, J., Goteti, G., and Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19(13):3088–3111.

Stevens, A. (2017). Temperature, wages, and agricultural labor productivity. mimeo.

World Bank (2008). *Public sector reform: What works and why? An IEG evaluation of World Bank support.* The World Bank.

Zamboni, Y. and Litschig, S. (2018). Audit risk and rent extraction: Evidence from a randomized evaluation in Brazil. *Journal of Development Economics*, 134:133–149.

A Construction of weather-related variables

Weather information in this study comes from the Princeton Meteorological Forcing Dataset. This reanalysis dataset combines the climate model information and observational data from various sources such as weather stations and satellite observations. It provides weather information even in places where observational data is scarce. The Princeton Meteorological Forcing Dataset is a 3-hourly dataset: weather variables are recorded at 0 am, 3 am, . . . , 9 pm at Greenwich time zone. Also, this is a gridded dataset at a resolution of 0.25×0.25 degrees. For details on the dataset, see Sheffield et al. (2006).

We use dry-bulb temperature, specific humidity, air pressure, and rainfall information in the dataset. To obtain each of these variables at each municipality, we use weather measures at four grid points surrounding municipality centroids and take the average of them, weighted by the inverse distance between the centroids and each of the four grid points. As daily weather measures, for variables other than the rainfall measure, we take the average values recorded between 9 am and 5 pm at the local time, accounting for the time zones and the daylight saving time. For instance, in a municipality in a state Acre, whose time zone is UTC−05:00, we use variables at 10 am, 1 pm, and 4 pm since they correspond to 3 pm, 6 pm, and 9 pm at Greenwich time zone, at which times weather information is recorded in the Princeton Meteorological Forcing Dataset. We take the average of these values and use it in our analyses as the temperature measurement on a particular day. Given these numbers, we take the average of them during fieldwork, which we use in the analyses. Since rainfall is recorded only daily, we take the average rainfall during the audit fieldwork as a rainfall measure in the regressions.

In the analyses, we use two different measures for temperature: dry-bulb temperature and wet-bulb temperature. Dry-bulb temperature is directly obtained from the Princeton Meteorological Forcing Dataset, and the wet-bulb temperature is calculated based on dry-bulb temperature, specific humidity, and air pressure, using the following formula (Geruso and Spears, 2018):

$$
\begin{aligned}
T_{wb} = T_{db} * \left[atan(0.151977 * (R + 8.313658)^{1/2}\right] + atan(T_{db} + R) \\
- atan(R - 1.676331) + 0.00391838 R^{3/2} * atan(0.023101 R) - 4.686035 \\
R = 0.263 * p * s * \left[\exp\left(\frac{17.67 T_{db}}{T_{db} + 243.5}\right)\right]^{-1},
\end{aligned}
$$

where T_{wb} is wet-bulb temperature (°C), T_{db} is dry-bulb temperature (°C), R is relative humidity (%), p is air pressure (Pa), and s is specific humidity.

B Appendix tables

Table B.1: Regression: Corruption reports and temperature (standard errors not wild bootstrapped)

VARIABLES	(1) Broad	(2) Narrow	(3) Broad	(4) Narrow
Wet bulb temp. (°C)	0.017*	0.021**	0.021**	0.023**
	(0.008)	(0.010)	(0.010)	(0.010)
Wet bulb temp. (one-year lag) (°C)			-0.011	-0.004
			(0.009)	(0.010)
Wet bulb temp. (one-year lead) (°C)			-0.001	-0.001
			(0.010)	(0.014)
Rainfall (mm/day)	0.000	0.001	0.001	0.001
	(0.005)	(0.005)	(0.005)	(0.005)
Observations	966	966	966	966
R-squared	0.27	0.32	0.27	0.32
Mesoregion FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean of outcome	0.736	0.439	0.736	0.439
SD of temperature in year t	3.550	3.550	3.550	3.550

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: The dependent variable in columns (1) and (3) is an indicator for broad corruption reported, and the dependent variable in columns (2) and (4) is an indicator for narrow corruption reported. Control variables include the share of the urban population, income per capita, log of population, population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Numbers in parentheses are the state-level clustered standard errors. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Table B.2: Regression: corruption reports and dry-bulb temperature

	(1) Broad	(2) Narrow	(3) Broad	(4) Narrow
Dry bulb temp. (°C)	0.008	0.012	0.008	0.014
	[-0.007, 0.021]	[-0.002, 0.025]	[-0.009, 0.024]	[0.001, 0.027]*
Dry bulb temp. (one-year lag) (°C)			-0.003	-0.002
			[-0.016, 0.011]	[-0.018, 0.015]
Dry bulb temp. (one-year lead) (°C)			0.001	-0.003
			[-0.019, 0.022]	[-0.024, 0.017]
Rainfall (mm/day)	0.003	0.005	0.003	0.005
	[-0.006, 0.014]	[-0.003, 0.015]	[-0.007, 0.014]	[-0.003, 0.015]
Observations	966	966	966	966
R-squared	0.26	0.32	0.26	0.32
Mesoregion FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean of outcome	0.736	0.439	0.736	0.439
SD of temperature in year t	4.275	4.275	4.275	4.275

Notes: The dependent variable in columns (1) and (3) is an indicator for broad corruption reported, and the dependent variable in columns (2) and (4) is an indicator for narrow corruption reported. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Numbers in square brackets are 90% confidence intervals, which are calculated with clustered wild-bootstrap at the state level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Table B.3: Regression: corruption reports (falsification tests)

	(1) Broad	(2) Broad	(3) Narrow	(4) Narrow
Wet bulb temp. (one-year lag) (°C)	-0.006		0.002	
	[-0.020, 0.009]		[-0.013, 0.018]	
Wet bulb temp. (one-year lead) (°C)		0.002		0.006
		[-0.012, 0.014]		[-0.017, 0.027]
Observations	966	966	966	966
R-squared	0.26	0.26	0.32	0.32
Mesoregion FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean of outcome	0.736	0.736	0.439	0.439
SD of temperature	3.387	3.722	3.387	3.722

Notes: The dependent variable in columns (1) and (2) is an indicator for broad corruption reported, and the dependent variable in columns (3) and (4) is an indicator for narrow reported corruption. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Numbers in square brackets are 90% confidence intervals, which are calculated with clustered wild-bootstrap at the state level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Table B.4: Regression: numbers of images and tables and temperature

	(1) Images	(2) Tables	(3) Images	(4) Tables
Wet bulb temp. (°C)	0.092	-0.305	0.144	-0.597
	[-0.679, 0.706]	[-1.179, 0.354]	[-0.600, 0.725]	[-1.533, 0.150]
Wet bulb temp. (one-year lag) (°C)			0.265	0.306
			[-0.508, 1.273]	[-0.231, 0.906]
Wet bulb temp. (one-year lead) (°C)			-0.273	0.391
			[-1.215, 0.412]	[-0.396, 1.097]
Rainfall (mm/day)	0.061	0.144	0.066	0.119
	[-0.262, 0.409]	[-0.177, 0.477]	[-0.255, 0.427]	[-0.212, 0.462]
Observations	966	966	966	966
R-squared	0.48	0.50	0.48	0.50
Mesoregion FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean of outcome	19.366	25.526	19.366	25.526
SD of temperature in year t	3.550	3.550	3.550	3.550

Notes: The dependent variable in columns (1) and (3) is the number of images in reports, and the dependent variable in columns (2) and (4) is the number of tables in the reports. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Numbers in square brackets are 90% confidence intervals, calculated with clustered wild-bootstrap at the state level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Table B.5: Regression: corruption reports and temperature (with state fixed effects)

	(1) Broad	(2) Narrow	(3) Broad	(4) Narrow
Wet bulb temp. (°C)	0.021	0.022	0.023	0.022
	[0.005, 0.034]**	[0.008, 0.036]**	[0.005, 0.039]**	[0.010, 0.036]***
Wet bulb temp. (one-year lag) (°C)			-0.010	-0.005
			[-0.025, 0.004]	[-0.017, 0.006]
Wet bulb temp. (one-year lead) (°C)			0.002	0.002
			[-0.012, 0.016]	[-0.020, 0.023]
Rainfall (mm/day)	0.000	0.002	0.000	0.002
	[-0.010, 0.012]	[-0.005, 0.012]	[-0.010, 0.013]	[-0.005, 0.012]
Observations	966	966	966	966
R-squared	0.15	0.21	0.15	0.21
State FE	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean of outcome	0.736	0.439	0.736	0.439
SD of temperature in year t	3.550	3.550	3.550	3.550

Notes: The dependent variable in columns (1) and (3) is an indicator for broad corruption reported, and the dependent variable in columns (2) and (4) is an indicator for narrow corruption reported. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Numbers in square brackets are 90% confidence intervals, calculated with clustered wild-bootstrap at the state level. *: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

C Appendix figures

Figure C.1: Distributions of temperature during fieldwork

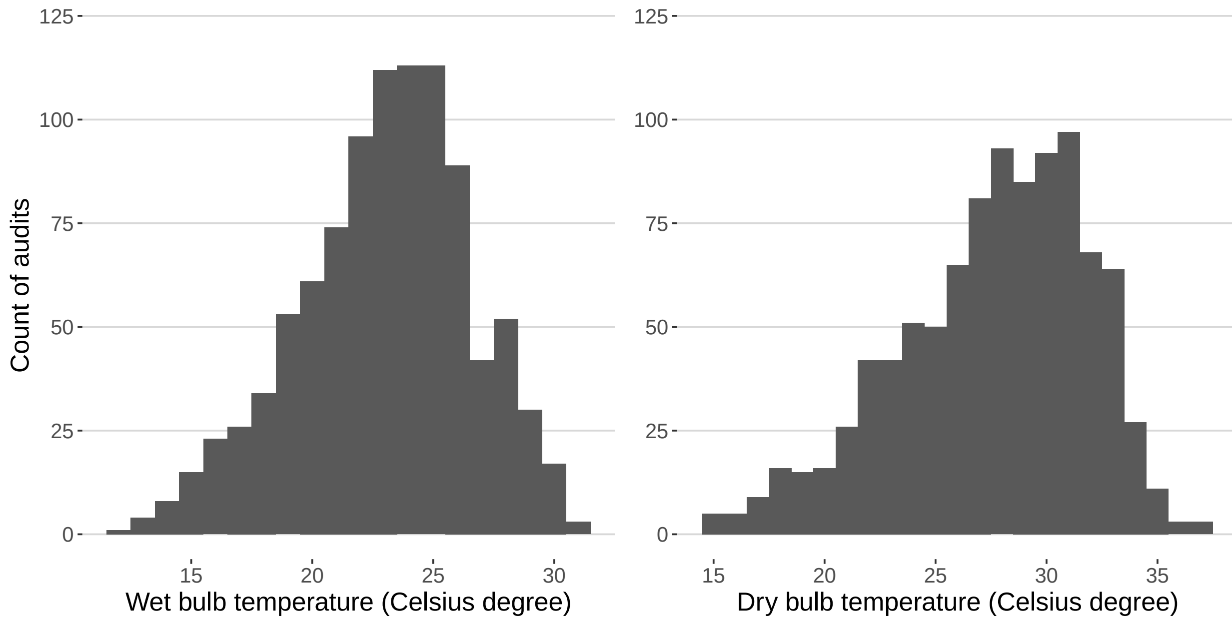
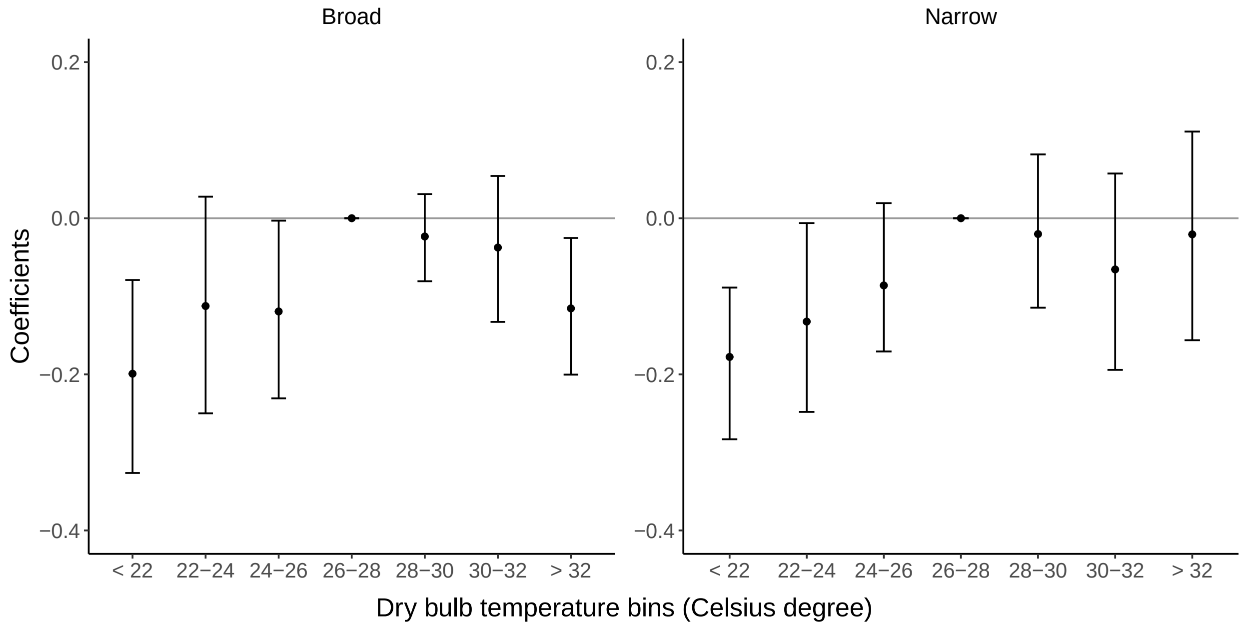
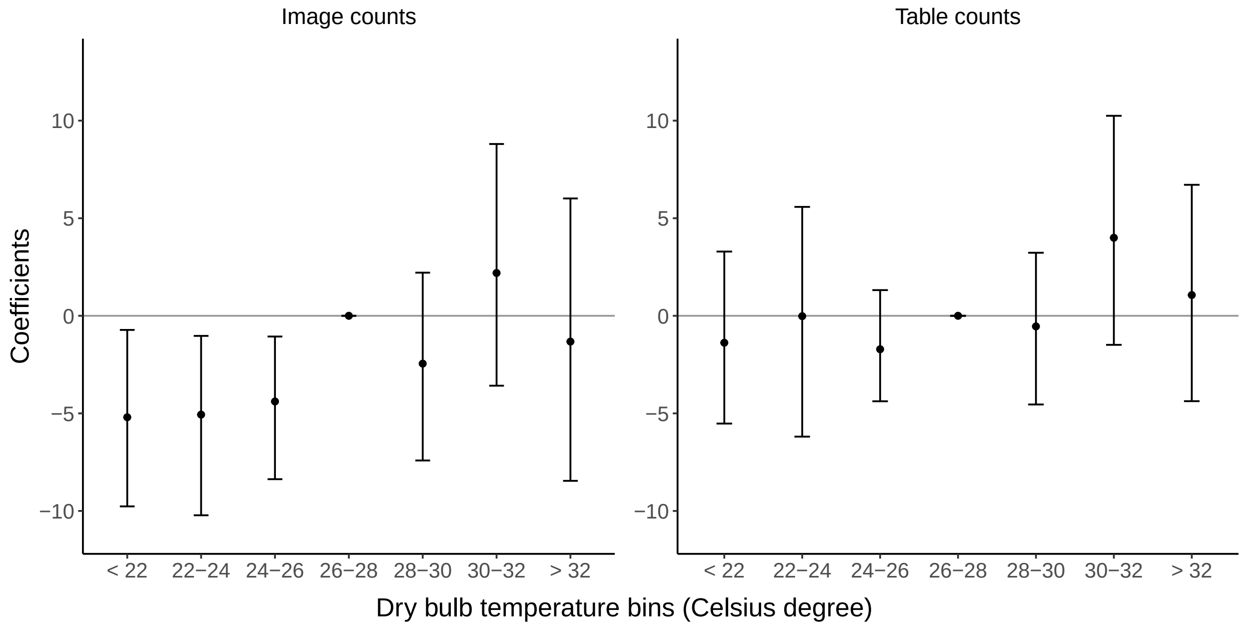


Figure C.2: Dry-bulb temperature effects on reported corruption



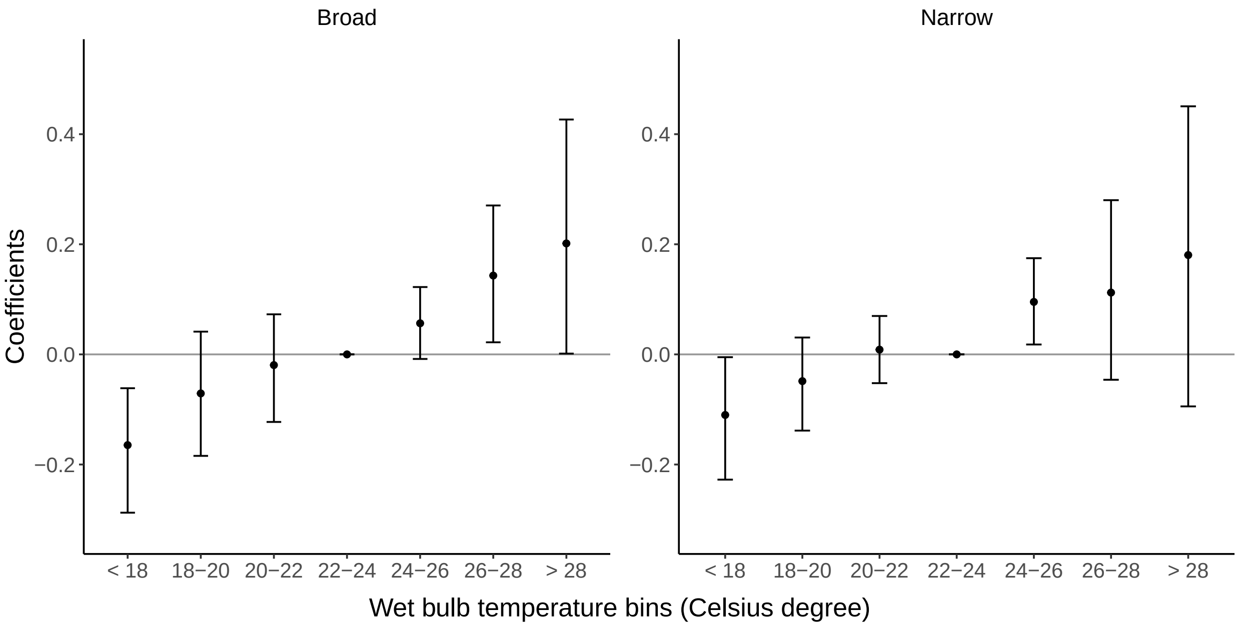
Notes: The figure shows the regression coefficients of an indicator for corruption (left: broad, right: narrow) reported on the dry-bulb temperature bins. Each temperature interval does not include the right end. The 90% confidence intervals are shown. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Confidence intervals are calculated with clustered wild-bootstrap at the state level.

Figure C.3: Dry-bulb temperature effects on numbers of images and tables



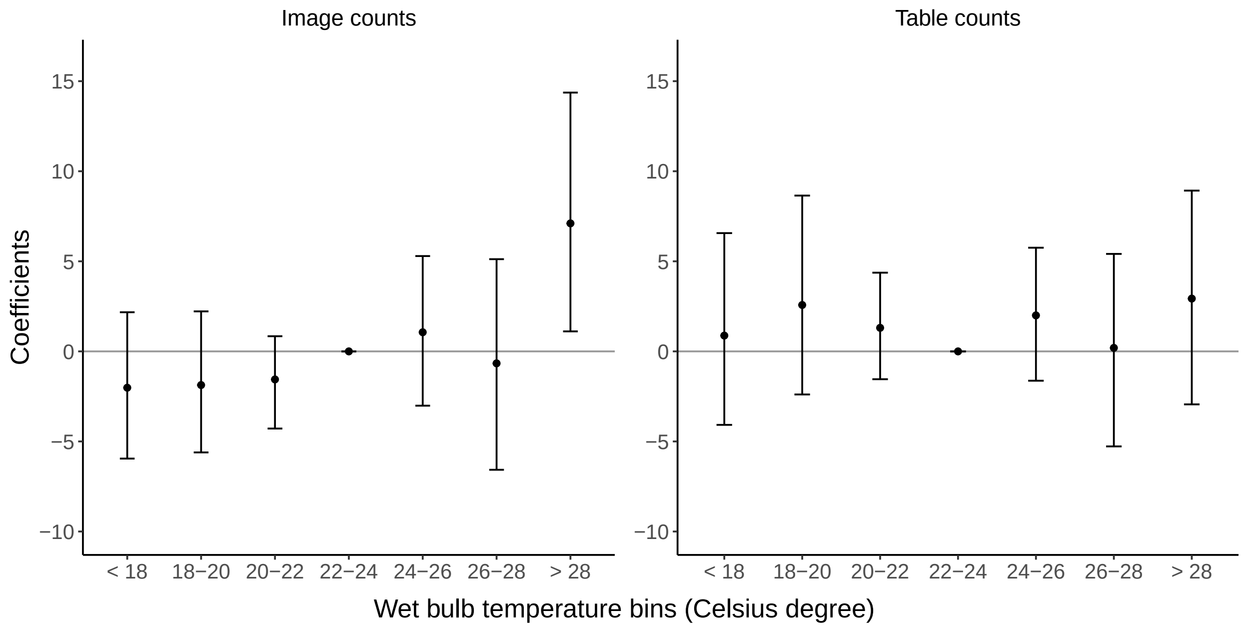
Notes: The figure shows the regression coefficients of the number of images (left) and the number of tables (right) on the dry-bulb temperature bins. Each temperature interval does not include the right end. The 90% confidence intervals are shown. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Confidence intervals are calculated with clustered wild-bootstrap at the state level.

Figure C.4: Wet-bulb temperature effects on reported corruption (with state fixed effects)



Notes: The figure shows the regression coefficients of an indicator for corruption (left: broad, right: narrow) reported on the wet-bulb temperature bins. Each temperature interval does not include the right end. The 90% confidence intervals are shown. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Audit wave and state fixed effects are included in the regressions. Confidence intervals are calculated with clustered wild-bootstrap at the state level.

Figure C.5: Wet-bulb temperature effects on number of images and tables (with state fixed effects)



Notes: The figure shows the regression coefficients of the number of images (left) and the number of tables (right) on the wet-bulb temperature bins. Each temperature interval does not include the right end. The 90% confidence intervals are shown. Control variables include the share of the urban population, income per capita, log of population, the population share below the poverty line, an indicator for second-time audit, and long-run (1980-2016) temperature and rainfall measures. Audit wave and state fixed effects are included in the regressions. Confidence intervals are calculated with clustered wild-bootstrap at the state level.

D Images and tables in corruption reports

To show that figures and tables are used in reports as evidence of corruption, we present examples of figures and tables in audit reports. Although only a few illustrations are provided, most figures and tables in other reports are similarly used.

In 2004, Estância, a municipality in Sergipe, has sent an invitation for a bid to three companies for the National School Nourishment Program (PNAE). The audit report pointed out that the three companies' associates are in the same family group, which demolishes the competitive nature of the procurement process. Moreover, the winning company was established just one month before the opening of the biddings. These suggest that the procurement practice was illegal. As supporting evidence, the report provides a table showing the information on invited companies (Figure D.1). The photos of envelopes sent to these companies are taken and provided as evidence in the report (Figure D.2).

Figure D.1: Example table from an audit report of Estância, Sergipe

N.	Empresa	Data de constituição	CNPJ	Quadro Societário
1	[redacted]	23/01/2004 (*)	[redacted]	[redacted]
2	[redacted]	14/06/1996	[redacted]	[redacted]
3	[redacted]	30/10/2000	[redacted]	[redacted]

(*) Observe-se que a empresa [redacted] – empresa vencedora – foi constituída cerca de 1 mês antes da abertura das propostas.

Notes: The table from an audit report of Estância, Sergipe, audited at the 10th lottery wave, is shown. The table shows the names, dates of establishment, identification numbers of firms (CNPJ), and structures of companies invited for the bidding. The table footnote emphasizes that the winning company was established one month before the bidding opening. Privately identifiable information is hidden, although, in the original report, the information is public and available at the CGU website.

Figure D.2: Example figures from an audit report of Estância, Sergipe

Notes: The photos from an audit report of Estância, Sergipe, audited at the 10th lottery wave, are shown. They are photos of envelopes sent to companies invited to the bidding. Privately identifiable information is hidden, although in the original report, the information is public and available at the CGU website.

Figures and tables are also used to provide evidence of mismanagement, besides severe corruption. Figure D.3 shows the photos in which equipment and insecticides are stored improperly. Figure D.4 is an example of a table listing inconsistencies in the registry of beneficiaries of the Bolsa Escola, a social assistance program targeted at the poor, now integrated into the Bolsa Família program.

Figure D.3: Example figures from an audit report of Estância, Sergipe

Notes: Figures from an audit report of Estância, Sergipe, audited at the 10th lottery wave, are shown. The left-top figure shows the storage of equipment for dengue fever, which is inappropriately placed in a humid environment. The right-top figure shows the wet floor next to the shelf in the left-top figure. The left-bottom figure shows the stocked equipment, where the near roof is wet. The right-bottom figure shows equipment that is not maintained correctly.

Figure D.4: Example table from an audit report of São Ludgero, Santa Catarina

3.3) Cadastro dos beneficiários do Bolsa Escola desatualizado e com diversos tipos de inconsistências.

Fato:

Verificamos que o cadastro dos beneficiários do Bolsa Escola no município apresenta uma série de inconsistências e está desatualizado, o que dificulta o acompanhamento e o controle do programa. As inconsistências identificadas foram as seguintes:

NIS do Responsável	Inconsistências Encontradas no trabalho de campo:
[redacted]	Aluno estuda na EEB São Ludgero e não na CEI Dom Gregório Warling.
[redacted]	Aluno estuda na EEB São Ludgero e não na CEI Dom Gregório Warling.
[redacted]	**a)** Aluno estuda na EEB São Ludgero e não na CEI Dom Gregório Warling. **b)** Data de nascimento registrada na escola é [redacted] e não [redacted].
[redacted]	Alunos estudam na EEB São Ludgero e não na CEI Dom Gregório Warling.
[redacted]	**a)** Família não reside mais no município. **b)** Nome da criança é [redacted] e não [redacted] .
[redacted]	Aluno estuda na EEB São Ludgero e não na CE Prof Henrique Buss.
[redacted]	**a)** Família não reside mais no município. **b)** Nome da criança registrado na escola é [redacted] e não [redacted].
[redacted]	Aluno estuda na EEB São Ludgero e não na CE Prof Henrique Buss.
[redacted]	Alunos estudam na EEB São Ludgero e não na CE Prof Henrique Buss.
[redacted]	Secretária da CE Prof Henrique Buss informou que aluna não estuda na escola. Aluno não identificado em nenhuma escola.
[redacted]	Data de nascimento registrada na escola é [redacted] e não [redacted]
[redacted]	Alunos estudam na EEB São Ludgero e não na EM Divina Providência. Data de nascimento de um dos filhos que está registrada na escola é [redacted] e não [redacted] Responsável possui dois números de NIS ([redacted] e [redacted].
[redacted]	Cadastro da escola do aluno no banco de dados do Bolsa Escola consta "crianças não identificadas". Verificamos que o mesmo estuda na EEB Bom Retiro.
[redacted]	Cadastro da escola do aluno no banco de dados do Bolsa Escola consta "crianças não identificadas". Prefeitura não informou em que escola do município criança estaria estudando.
[redacted]	Cadastro da escola do aluno no banco de dados do Bolsa Escola consta "crianças não identificadas". Recebemos a informação que o aluno estuda em um supletivo da cidade e não no ensino fundamental de São Ludgero.
[redacted]	Prefeitura não informou em que escola do município as crianças de NIS [redacted] e [redacted] estariam estudando.

Notes: A table from an audit report of São Ludgero, Santa Catarina, audited at the 10th lottery wave, is shown. It presents information on the registration ID of Bolsa Escola beneficiaries and inconsistencies in their registrations. Privately identifiable information is hidden, although, in the original report, the information is public and available at the CGU website.