Online Resource 1 (ESM_1)

Beyond the Ping-Pong Table: Verifiable Commitment and the Governance of Employee Well-Being Spending

Journal of Business Economics

Authors: withheld for double-anonymous review

Contents: S1 The signaling model. S2 Data and sample construction. S3 Locating the human capital section. S4 Coding rules. S5 Validation audits. S6 Additional results. S7 Reproduction.

S1 The signaling model

Setting. Time is discrete, $t = 0, 1, 2, \ldots$, and payoffs are discounted by $\delta \in (0, 1)$. A firm is either committed (type *C*) or uncommitted (type *U*); its type is private information. At $t = 0$ the firm decides whether to announce a well-being programme. Maintaining the programme costs $c > 0$ per period. Outsiders (employees, applicants, investors) grant a premium $\pi > 0$ in every period in which they believe the firm is committed. The programme's verifiability determines the probability $p \in [0, 1]$ that a withdrawal is observed by outsiders. The three BCR conditions are interpreted as raising p.

Committed type. The committed firm derives an intrinsic per-period value g from the programme (productivity, retention, owner preferences) and never withdraws it. It announces if $(\pi + g - c)/(1 - \delta) \geq 0$, which holds whenever $g + \pi \geq c$. We assume this participation constraint holds.

Uncommitted type. The uncommitted firm values only the premium. In each period a downturn occurs with probability q; at the first downturn the uncommitted firm withdraws the programme. If the withdrawal is observed (probability p), outsiders revise their beliefs, the premium ends, and the firm incurs an additional loss $R > 0$ that reflects the cost of a broken promise relative to no promise (breach of the psychological contract, loss of trust among investors). If the withdrawal is not observed (probability $1 - p$), the firm continues to receive the premium without paying for the programme; we write the value of this continuation as $\Pi = \pi/(1 - \delta)$.

Value of imitation. Let W be the uncommitted firm's value of announcing at the start of a period in which the programme is maintained. Then

$$W = (\pi - c) + \delta[(1 - q)W + q((1 - p)\Pi - pR)],$$

so that

$$W\,[1 - \delta(1 - q)] = (\pi - c) + \delta q[(1 - p)\Pi - pR].$$

Since $1 - \delta(1 - q) > 0$, W has the sign of the right-hand side, which gives condition (1) in the main text. Not announcing yields zero, because outsiders then infer that the firm is uncommitted. Imitation is unprofitable, and the announcement separates the types, when the right-hand side is non-positive, which is condition (2).

Threshold verifiability. Solving condition (2) for p gives the minimum verifiability at which an announcement separates:

$p^* = (\pi + \delta q \Pi - c) / (\delta q (R + \Pi))$.

If $p^* \leq 0$ the programme is costly enough to separate without verification (Spence's case). If $p^* > 1$ no degree of verifiability suffices, which happens when the premium is large relative to the penalty for an observed withdrawal.

Comparative statics. For an interior threshold:

- $\partial p^*/\partial R < 0$: the larger the loss from a broken promise, the less verification is needed.
- $\partial p^*/\partial c < 0$: more costly programmes need less verification; cost and verifiability are substitutes.
- $\partial p^*/\partial \pi > 0$: a larger premium for appearing committed raises the temptation to imitate and therefore the verification needed.
- $\partial p^*/\partial q < 0$ whenever $\pi > c$: in volatile environments the expected penalty for silent exit is realised more often, so a given level of verifiability does more work. Writing $p^* = (\pi - c)/(q(\delta R + \delta \Pi)) + \Pi/(R + \Pi)$ shows the sign directly.

Interpretation. The model is deliberately minimal. It abstracts from partial withdrawal, from learning across periods, and from heterogeneity among receivers. Its purpose is to show that, for a firm that signals repeatedly, the observability of withdrawal can substitute for the cost of the signal, and that unverifiable spending cannot separate the types unless its cost exceeds the premium together with the value of a silent exit. The propositions in the main text follow from this logic; the model does not by itself establish any of them empirically.

S2 Data and sample construction

Table S1 Sample construction

Step	Firms	Filings or firm-years
Firms in the S&P 500 constituent list (October 2026), unique CIK numbers	500	
Form 10-K filings on EDGAR filed 1 January 2021 to 27 October 2026	497	2,878 filings
One filing per firm and fiscal year, fiscal years 2020 to 2025	496	2,828 firm-years
Less: section shorter than 40 words or incorporated by reference		134
Less: no human capital section located		65
Coded sample	473	2,629 firm-years
Long-window sample (firm observed in fiscal 2020 or 2021 and in fiscal 2024 or 2025)	440	
Of which: firms with a well-being claim at the start	319	
Of which: with financial data for both years (Table 3 sample)	293	

Fiscal years ending January to March are assigned to the previous calendar year, following the convention in retail. When a firm filed two 10-Ks for the same fiscal year, the later filing is kept. Amended filings (10-K/A) are not used. The constituent list reflects index membership in October 2026, so firms that left the index between 2021 and 2026 are not in the sample.

S3 Locating the human capital section

Each filing's primary document was converted to text by replacing block-level HTML tags and table cells with line breaks and removing the remaining tags, hidden inline XBRL headers, scripts, and styles. Running page headers (for example "Table of Contents", page numbers, and repeated "Item 1.

Business” lines) were removed. Item 1 was identified as the occurrence of the “Item 1. Business” heading followed by the largest block of text before Item 1A, which skips the table of contents.

Within Item 1, the section begins at the first line of at most 80 characters that starts with “Human capital” (optionally “Our human capital”, “Human capital resources”, or “Human capital management”). If none exists, the parser looks, in this order, for a line that consists only of “Employees”, “Our people”, “People”, “Workforce”, “Human resources”, “Talent”, “Associates”, “Team members”, “People and culture”, or “Employees and human capital”; for any short heading that contains the words “human capital” (for example “Team members and human capital management”); for a paragraph that opens with a run-in heading such as “Employees:” or “Human capital.”; and finally for the first line within Item 1 that mentions human capital. A match in the second half of the filing is rejected, because it indicates that the Item 1 heading was not recognised.

The section ends at the first of three kinds of line: (a) a line of at most 90 characters that starts with Item 1A or a standard Item 1 heading on another topic (available information, executive officers, intellectual property, regulation, environmental matters, seasonality, working capital, properties, competition, research and development, and similar); (b) a line of at most 70 characters, without final punctuation, that starts a standard non-workforce topic (sustainability, environmental, climate, segments, products, customers, markets, insurance, risk management, suppliers, patents, cybersecurity, and similar) and is followed by a line of at least 40 characters; or (c) a heading-like line of at most 70 characters without workforce vocabulary, where the following 700 characters include a prose line of at least 100 characters and also contain no workforce vocabulary. Sections shorter than 40 words are treated as incorporated by reference. The exact rules and vocabularies are in the code (Online Resource 3, file 01_download_filings.py).

S4 Coding rules

Sections were split into sentences at full stops, question marks, and exclamation marks that end a line or are followed by an upper-case letter or an opening quotation mark; zero-width characters were removed first. All dictionaries are case-insensitive regular expressions. The full list of terms and the definition of every variable are in the codebook (Online Resource 2, sheet “Codebook”). Monetary amounts and usage figures count only if they occur within 150 characters of a well-being term in the same sentence, so that unrelated numbers in long sentences are not attributed to well-being claims. Usage figures in sentences about engagement surveys, participation rates in surveys, resource networks or groups, stock plans, learning or training hours, volunteering, community service, philanthropy, or remote work are excluded, because they do not describe the use of a well-being programme. Reduction statements require a verb of reduction (reduce, eliminate, discontinue, suspend, cut, freeze, terminate, end, phase out, scale back) followed within four words by an employee benefit, perk, leave, 401(k) or matching contribution, wellness or well-being programme, stipend, subsidy,

salary, wage, or bonus; sentences about injuries, risks, incidents, bias, waiting periods, store hours, employee costs or premiums, decisions not to act, or the absence of reductions (“no salary reductions”, “did not reduce”) are excluded, and “end” is not counted in “year-end” or “end of the year”. Six firm-years, from three firms, contain a reduction statement under this rule; all six were read and all describe an actual reduction.

S5 Validation audits

Sentence audit. Forty sentences coded as well-being claims were drawn at random (seed 7) and read in full. A sentence was judged correct if it refers to the firm’s attention to, or programmes for, the well-being of its own employees. Thirty-eight of the 40 (95 percent) are correct. The two false positives are an executive officer biography that mentions a financial planning function and a sentence about counselling that the human resources team gives to leaders, managers and employees in the course of its work. The sample and the judgements are in data/audit/audit_sentences.csv (Online Resource 3).

Figure audit. All 41 firm-years coded as making a claim with a figure were read. In 25 firm-years, from 12 firms, the figure belongs to the well-being programme or support itself: per-employee amounts (15), rates of use or pay during the pandemic (5), firm-level pandemic spending totals or relief funds (4), and a wage floor presented as supporting financial well-being (1). In the other 16, the figure belongs to another benefit listed in the same sentence (9), concerns a community programme (5), describes availability rather than use (1), or is a wage increase whose only link to well-being is a section heading attached to the sentence (1). The list and the judgements are in data/audit/audit_figures.csv.

Section audit. Twenty sections drawn at random (seed 11) were used while the parser was developed; after the last change all of them start and end at the section boundary, but because the parser was revised on this sample it is not a test of accuracy. A second, independent sample of twenty sections (seed 23, excluding the first sample) was drawn after the parser was final and compared with the full text of each filing. All of them start at the human capital heading, and 17 end at the boundary of the section. In the other three the extract runs on into the following section (insurance reserves and investments; regulatory and environmental matters; a sustainability paragraph). Overruns of this kind lengthen a section but do not remove well-being text. Both samples and the judgements are in data/audit/.

Truncation check. For every filing in which a section was located (2,811), the text between the end of the extract and the next standard heading (or Item 1A) was searched for well-being terms. Such terms occur in 46 filings from 17 firms. None is an employee well-being statement that belongs under the human capital heading: 43 occur in a separate environmental, social and governance, sustainability, climate, or governance section that the firm placed after its human capital section, 2 in

product descriptions, and 1 in a description of community programmes (data/audit/audit_truncation.csv).

Drops. Each of the 22 full-window drops was checked against the full text of the end-of-window filing. In 21 the filing contains no well-being reference in or around its human capital section. In one case (a railway equipment manufacturer) the well-being sentence moved from the human capital section to a compensation and benefits paragraph that the firm placed before its human capital heading (data/audit/audit_drops.csv).

S6 Additional results

Table S2 Year-to-year changes in well-being claims

Fiscal years	Firm pairs	Claim dropped (% of firms with a claim)	Claim added (% of firms without a claim)
2020 to 2021	361	2.3	20.6
2021 to 2022	442	4.2	15.0
2022 to 2023	448	0.6	10.2
2023 to 2024	449	2.6	8.2
2024 to 2025	453	1.7	5.0
All	2,153	2.3	11.8

A firm pair is a firm observed in two consecutive fiscal years. Over all pairs there are 37 drops and 60 additions; none of the drops occurs in a filing whose human capital section contains a reduction statement.

Table S3 Robustness of the determinants of dropping a well-being claim

Specification	N	Change in log section length	Well-being sentences at start	Change in return on assets	Log total assets
Baseline specification	293	−0.158***	−0.018**	0.087	−0.014
		(0.044)	(0.008)	(0.133)	(0.010)
Narrow well-being dictionary	189	−0.122	−0.024*	0.714*	0.042
		(0.076)	(0.014)	(0.412)	(0.027)
Claims excluding COVID-19 sentences	283	−0.158***	−0.016**	0.102	−0.015
		(0.044)	(0.007)	(0.134)	(0.011)
Baseline fiscal year 2022	310	−0.158***	−0.005	0.018	−0.012
		(0.047)	(0.004)	(0.100)	(0.008)
Heading-located sections only	290	−0.160***	−0.018**	0.079	−0.013
		(0.045)	(0.008)	(0.134)	(0.011)
Sections of at least 100 words in both years	291	−0.123***	−0.018**	0.084	−0.013
		(0.042)	(0.007)	(0.125)	(0.010)
Larger firms (assets above median)	143	−0.116*	−0.020*	−0.073	−0.018
		(0.069)	(0.012)	(0.233)	(0.014)
Smaller firms (assets below median)	150	−0.207***	−0.018	0.258	0.016
		(0.067)	(0.012)	(0.187)	(0.037)
Excluding financial firms	246	−0.171***	−0.018**	0.114	−0.021
		(0.050)	(0.009)	(0.140)	(0.015)

All specifications are linear probability models with the controls and sector effects of Table 3, column 3, and heteroskedasticity-robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. In the narrow-dictionary and COVID-excluded rows, both the starting claim and the outcome are measured with the alternative definition. With fiscal 2022 as the starting year, 338 firms have a claim at the start and 3.6 percent drop it by the end.

Table S4 Persistence of well-being claims by whether the initial claim carried a figure

Initial claim	Firms	Claim kept at end of window	Share kept (%)
Coded as carrying a figure	7	6	85.7
Not coded as carrying a figure	312	291	93.3

Fisher’s exact test, two-sided: p = 0.40. With 7 firms in the first row, the comparison cannot test Proposition 4.

Intensity outcome. Among all 407 firms in the long window with financial data, the change in the log of one plus the number of well-being sentences was regressed on the change in log section length, the change in return on assets, log total assets, and sector effects. The coefficient on section length is 0.437 (standard error 0.052, p < 0.001) and the coefficient on the change in return on assets is 0.040 (standard error 0.302, p = 0.90).

Section length. Between the start and the end of the long window, 58.9 percent of firms shortened their human capital section; the median change was −0.081 log points.

S7 Reproduction

Online Resource 3 (ESM_3.zip) contains the code, the constituent list, the XBRL frames, the extracted human capital sections, the audit files, and a README. Running bash code/run_all.sh reproduces every number in the paper and in this document from the included files (output in out/results.json); bash code/run_all.sh --download first re-downloads all filings from EDGAR and re-extracts the sections. Filings amended or removed on EDGAR after October 2026 could change a re-download slightly. The audit judgements in data/audit/ were made by reading the text; the code draws the samples and lists the cases, and python3 code/07_audits.py score computes the audit counts. The truncation check (code/08_truncation_check.py) needs the full filing texts, which the download step stores in data/text_cache/ (not included because of their size).