

Research Article

Visible but Unverified: How Public Signals Diverge from Formal Records in Israeli Digital Health Startups

Xiao-yuan Wang^{1,2}

1. School of Management and Economics, University of Electronic Science and Technology of China, China; 2. Center for Israeli Studies, University of Electronic Science and Technology of China, China

Publicly observed connections give outsiders concrete information about young digital health firms. We ask whether those connections predict two public evidence outcomes in the same way. The outcomes are medical news evidence about clinical or regulatory activity and formal records from clinical, regulatory, or research systems. We assemble a company-level dataset that connects digital health startups to company profiles, public news, named people and institutions, and formal records. We define a public connection as a publicly observed named person or institution that links a startup to another startup in the study network. Among clinically oriented firms with retrieved news, connected firms have 2.16 times higher odds of medical news evidence and a 14.1 percentage-point higher probability in the linear model. However, the effect of public connections on formal records is small, with an odds ratio of 1.12. Formal records align with company age and funding amount, with odds ratios of 2.41 and 1.87. These patterns show that public connections are more of a visibility measure than a measure of formal innovation records.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

1. Introduction

Evaluating digital health startups is difficult because clinical reliability, regulatory progress, and scientific credibility are costly to verify. Outsiders therefore rely on observable signals, including public connections to founders, advisers, investors, hospitals, universities, and research organizations. Building on research on information asymmetry and entrepreneurial signaling ^{[1][2][3][4][5][6][7]}, we ask whether these public connections predict concrete evidence of startup progress.

The first progress outcome is medical news evidence, which captures public text about clinical or regulatory activity, including pilots, hospital partnerships, clinical use, health-sector deployment, and regulatory steps. The second outcome is formal records, which capture entries in clinical-trial registries, device-regulatory databases, and research indexes.

We assemble a company-level dataset of Israeli digital health startups. The dataset links each firm to company characteristics, retrieved public articles, public connections to named people and institutions, and formal records from clinical, regulatory, and research systems. It contains 526 core firms, and the 330 firms with retrieved public articles form the main analysis sample.

The results show a clear split between medical news evidence and formal records. Among article-visible firms, public connections are associated with 2.16 times higher odds of medical news evidence. The corresponding formal-record estimate is small and imprecise. Formal records rise with company age and funding. Estimates from all 526 core firms preserve this pattern.

The contribution is to show that the predictive content of a public connection depends on the evidence system being studied. Prior work establishes affiliations, alliances, patents, team experience, and venture-capital reputation as signals of young-firm quality [\[2\]\[3\]\[4\]\[5\]\[7\]](#). We distinguish evidence created through public reporting from evidence created by clinical, regulatory, and scientific systems. The results show that these evidence sources reflect distinct processes and measure different dimensions of startup progress.

The paper proceeds in four steps. Section 2 develops the public-connection hypotheses. Section 3 describes the dataset and sample construction. Section 4 presents the logit model and linear probability checks. Section 5 presents the main estimates and robustness checks.

2. Conceptual Framework

Digital health startups must establish quality for several audiences. Investors assess growth potential, hospitals assess clinical reliability and workflow fit, regulators assess product claims, and research partners assess scientific credibility. These qualities are difficult to verify early in a firm's life. Public connections can reduce this information gap by linking a startup to named people and institutions whose identities, experience, and participation are publicly observable.

We evaluate public connections against two forms of evidence. Medical news evidence records whether retrieved articles describe clinical use, regulatory steps, hospital pilots, or health-sector deployment.

Formal records indicate whether the firm appears in a clinical-trial registry, device-regulatory database, or research index. The first outcome is produced through public reporting; the second requires entry into an external record system.

Public connections should increase the probability of medical news evidence because they place startups in networks that support both medical activity and its public communication. Hospitals and universities can provide users, clinical settings, and expert partners; investors and advisers can supply commercial contacts, regulatory knowledge, and communication channels. These resources help firms undertake pilots, deployments, and regulatory steps that become reportable events. Named counterparties also lend credibility and often have their own incentives to publicize joint activity. Entrepreneurial-finance studies similarly show that affiliations and alliances convey information about young firms [\[3\]\[4\]\[7\]](#).

Hypothesis 1. Among firms with retrieved news, public connections increase the probability of medical news evidence.

Public connections can also support the production of formal records. Hospitals and universities can supply study sites, investigators, scientific collaborators, and institutional credibility. Investors and advisers can supply capital and procedural knowledge for trials, regulatory submissions, and research publication. These resources lower the organizational cost of entering clinical, regulatory, and research systems.

Hypothesis 2. Public connections increase the probability of formal records.

The empirical test compares firms with at least one shared public connection to firms whose observed affiliations are unique or unshared within the study network. The estimates ask whether connected firms are more likely to have each outcome after accounting for firm maturity, resources, and technology area.

3. Data

Our dataset contains 526 Israeli digital health startups whose primary products serve human health care, drawn from Startup Nation Finder.¹ It links company profiles to retrieved public articles, named people and institutions, and formal records [\[8\]](#). The database was frozen on June 26, 2026, and the article collection runs through June 13, 2026. Among these firms, 330 have at least one retrieved public article and form the main analysis sample. This restriction ensures that a zero for medical news evidence means the retrieved text contains no qualifying medical content. The full core sample is used to test whether the

results extend to firms without retrieved articles. Table 1 reports the size, role, and key-variable counts for the main analysis sample and the full-core robustness sample.

Sample	Role	Firms	Public connections	Medical news evidence	Formal records
Visible core sample	Main analysis	330	146	221	107
Full core sample	Robustness analysis	526	168	228	127

Table 1. Analysis samples and key-variable counts

Notes: The full core sample contains 526 Israeli digital health startups whose primary products serve human health care. The visible core sample contains the 330 firms with at least one retrieved public article. The final three columns count firms for which PublicConnection, MedicalNewsEvidence, and FormalRecord equal one.

Table 2 defines and summarizes the model variables for the 330 firms in the main analysis sample. The variables are ordered as the two dependent variables, the main public-connection measure and its alternatives, and the controls.

Variable	Description	N	Mean	SD	Median
MedicalNewsEvidence	Indicator for article content on clinical or regulatory activity	330	0.67	0.47	1.00
FormalRecord	Indicator for a clinical, regulatory, or research-indexed record	330	0.32	0.47	0.00
PublicConnection	Indicator for sharing a named entity with another core firm	330	0.44	0.50	0.00
NeighborDegree	Number of other core firms reached through shared named entities	330	2.33	4.45	0.00
WeightedDegree	Sum of confidence- and frequency-adjusted connection weights	330	1.55	2.78	0.00
ConnectionPageRank	Weighted-network PageRank, multiplied by 1,000	330	2.40	2.72	0.68
ArticleCount	Number of retrieved public articles	330	3.63	5.68	2.00
CompanyAge	Years since founding	330	10.54	6.28	9.00
FundingAmount	Funding in million USD; zero when undisclosed	330	22.80	48.11	4.42
MissingFunding	Indicator for an undisclosed funding amount	330	0.28	0.45	0.00

Table 2. Summary statistics for the main analysis sample

Notes: All rows use the visible core sample. Statistics are reported in natural units. Means of indicators are sample shares. Regressions use $\log(1 + x)$ standardized within the estimation sample for NeighborDegree, WeightedDegree, ConnectionPageRank, ArticleCount, and FundingAmount; CompanyAge is standardized in levels. FundingAmount is set to zero when undisclosed and enters jointly with MissingFunding. Funding-stage and technology-family indicators are categorical controls and are omitted from this table.

The first dependent variable, MedicalNewsEvidence, equals one when a firm's retrieved articles describe clinical use, hospital pilots, health-sector deployment, or regulatory activity. The second, FormalRecord, equals one when the firm has an accepted match in a clinical-trial registry, device-regulatory database, or research index. These definitions keep the two outcomes distinct while applying them to the same firms.

Formal records are assembled from three public systems. ClinicalTrials.gov provides clinical-study registrations ^[9]. openFDA provides device-regulatory records, including 510(k) clearances ^[10]. OpenAlex provides research-indexed works ^[11].

We construct source-specific matches from identity fields. The relevant fields are sponsors and collaborators in ClinicalTrials.gov, applicants in openFDA, and affiliations and institutions in OpenAlex. Queries use standardized company names and a limited set of company aliases. A candidate enters FormalRecord only when the company name or alias appears in the relevant field. The replication files retain the query alias, source identifier, match reason, and source URL for each accepted record.

Public connections are constructed from a bipartite firm–entity network. One set of nodes contains the 526 core firms. The other contains named founders, team members, advisers, investors, hospitals, universities, research organizations, and corporate partners observed in public sources. For each firm, the network records whether any of these named people or institutions is shared with another core firm.

We project this bipartite network into a firm-to-firm network. Let S_i be the set of named entities linked to firm i . Two firms are connected when $S_i \cap S_j \neq \emptyset$. The main variable, $\text{PublicConnection}_i$, equals one when firm i shares at least one named person or institution with another core firm and zero otherwise. A zero is relative to this observed network: the firm's affiliations may be unique within the database, private, unobserved, or located outside the network. The indicator therefore measures whether a firm occupies a publicly observable shared path in the Israeli digital health ecosystem.

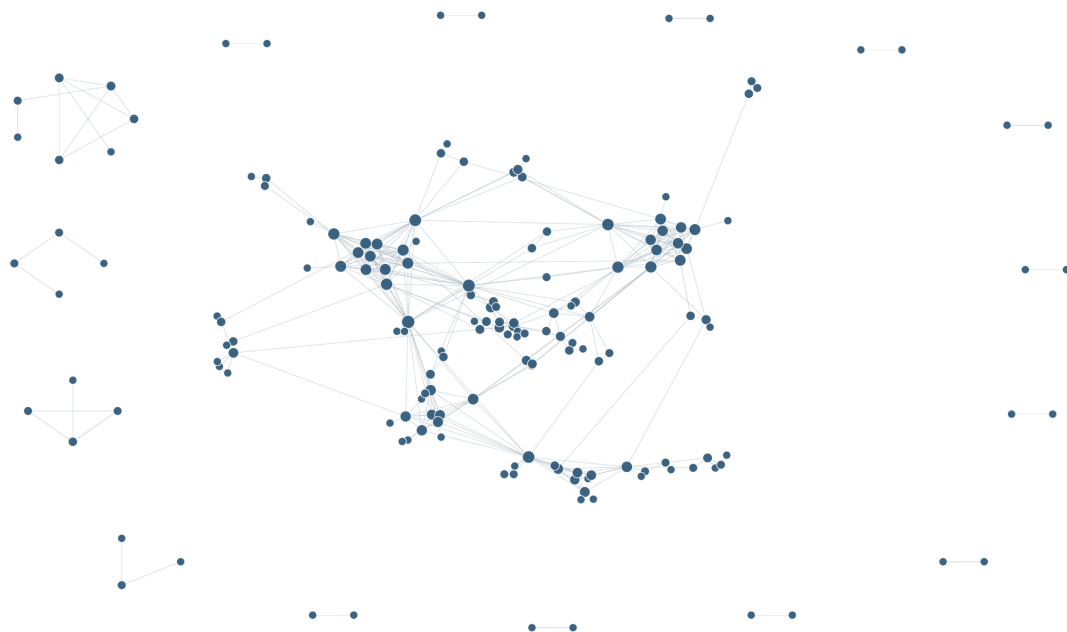


Figure 1. Public connections among connected startups. The figure shows the 168 core firms that share at least one named person or institution with another core firm. The graph contains 402 firm-to-firm edges. Node area increases with the number of other core firms reached, and edge width increases with the number of shared entities. Positions are chosen for readability.

Figure 1 displays the connected portion of the projected startup network. It maps the 168 core firms for which $\text{PublicConnection} = 1$ and the 402 edges among them. A large connected component contains 128 firms, while the remaining firms form smaller clusters through shared named people or institutions.

Three alternative measures describe the intensive margin of public connections. Neighbor degree counts the other core firms reached through at least one shared named entity. Weighted degree sums the projected edge weights attached to those neighbors. Each shared entity contributes the geometric mean of the two firm–entity evidence weights, discounted by the logarithm of the number of firms linked to that entity; institution contributions receive a factor of 0.9. Firm–entity weights combine source confidence and repeated evidence and are capped at 2.5. PageRank measures centrality in this weighted firm projection. Table 2 gives the corresponding variable names. The regressions transform each measure using $\log(1 + x)$ and standardize it within the estimation sample.

The remaining model variables account for observable differences in media coverage, firm maturity, financial resources, development stage, and product domain. ArticleCount measures the volume of retrieved text available to classify MedicalNewsEvidence. CompanyAge measures years since founding, and FundingAmount records total funding in millions of U.S. dollars. Undisclosed funding is set to zero and identified by MissingFunding. Funding-stage indicators distinguish pre-seed or pre-funding, seed, A/B early growth, C-plus late growth, and mature, public, or acquired firms. Technology-family indicators distinguish care management, clinical workflow, digital therapeutics, devices and wearables, drug and pharma tools, imaging and diagnostics, remote monitoring, and other technologies. Funding-stage and technology-family indicators allow baseline outcome rates to vary across development stages and product domains. Both outcome models use the same controls.

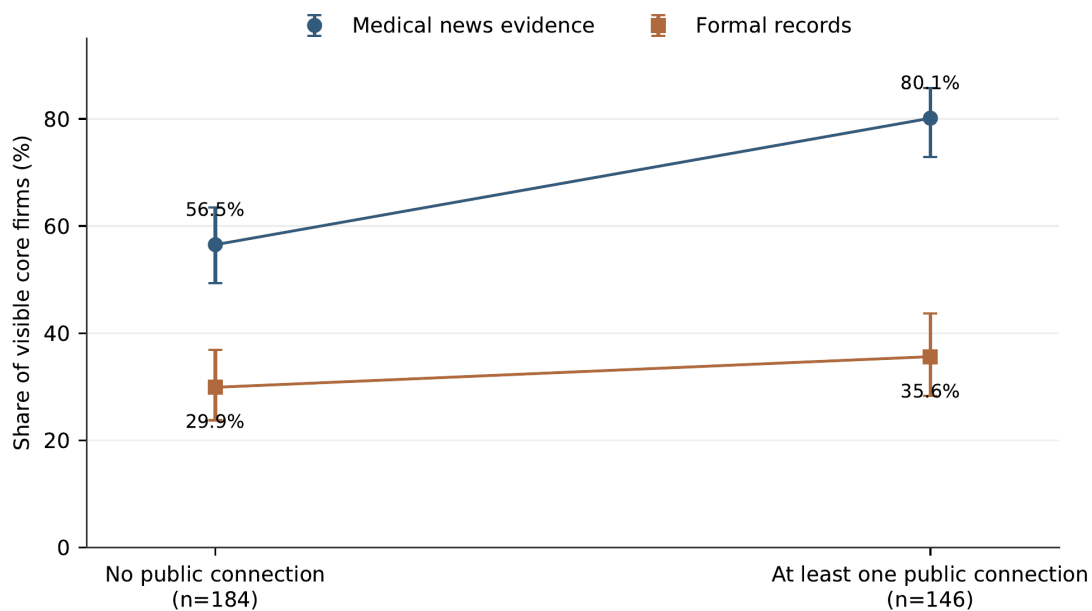


Figure 2. Public connections and evidence outcomes among visible firms

Figure 2 plots outcome rates by public-connection status in the visible core sample. Medical news evidence rises from 56.5 percent among firms without a public connection to 80.1 percent among firms with at least one connection. Formal records rise from 29.9 percent to 35.6 percent. The raw data therefore show a much larger connection gradient for medical news evidence.

4. Empirical Design

Our empirical strategy estimates how public connections relate separately to medical news evidence and formal records within the same firms. The main sample includes core startups with at least one retrieved public article, allowing medical news evidence to be classified from observed text for every firm.

We first estimate the MedicalNewsEvidence model

$$\begin{aligned} \Pr(\text{MedicalNewsEvidence}_i = 1) = \Lambda(&\alpha + \theta\text{PublicConnection}_i \\ &+ \rho\text{ArticleCount}_i + \eta\text{CompanyAge}_i \\ &+ \phi\text{FundingAmount}_i + \mu\text{MissingFunding}_i \\ &+ \gamma_{s(i)} + \delta_{f(i)}), \end{aligned} \quad (1)$$

where $\Lambda(\cdot)$ is the logistic function. The continuous covariates use the transformations reported in Table 2. The terms $\gamma_{s(i)}$ and $\delta_{f(i)}$ denote funding-stage and technology-family controls; A/B early growth and care management are the omitted categories.

We estimate the same specification separately with FormalRecord_i as the dependent variable and re-estimate every coefficient. The coefficient on $\text{PublicConnection}_i$ is the parameter of interest in each model. Hypothesis 1 predicts a positive coefficient in the medical-news model, and Hypothesis 2 predicts a positive coefficient in the formal-record model. This common specification holds the sample and conditioning variables constant across the two outcomes.

The main text reports logit odds ratios. The appendix re-estimates both outcome models as linear probability models with the same regressors, so their public-connection coefficients are expressed in percentage points.

To check the robustness of the main results, we use three complementary specifications. First, we estimate both outcome models on all 526 core firms and add AnyArticle_i to distinguish firms with and without retrieved text. Second, we replace the binary public-connection indicator with neighbor degree, weighted degree, and PageRank. These models test whether the results also appear along the intensive margin of network position. Third, we estimate linear probability counterparts for the visible and full core samples. The appendix also reports financing-news and partnership-news outcomes as extensions beyond the paper's two central outcomes (Tables A4 and A1).

5. Results

5.1. Main estimates among visible firms

Table 3 presents the main logit estimates for article-visible core firms. Both columns use the same sample and the same covariates.

	Medical news evidence	Formal records
PublicConnection	2.16*** [1.24, 3.77]	1.12 [0.60, 2.08]
ArticleCount	2.27*** [1.60, 3.20]	0.85 [0.60, 1.21]
CompanyAge	1.02 [0.73, 1.43]	2.41*** [1.49, 3.88]
FundingAmount	1.34 [0.78, 2.30]	1.87** [1.14, 3.07]
MissingFunding	1.01 [0.40, 2.52]	1.96 [0.69, 5.55]
Observations	330	330
Positive outcomes	221	107
McFadden pseudo R^2	0.179	0.216
Funding-stage controls	Yes	Yes
Technology-family controls	Yes	Yes

Table 3. Main logit estimates among article-visible firms

Notes: Cells report odds ratios with heteroskedasticity-robust 95 percent confidence intervals in brackets. *PublicConnection* is an indicator. *ArticleCount* and *FundingAmount* enter as $\log(1 + x)$ standardized within the estimation sample; *CompanyAge* is standardized in levels. *FundingAmount* is set to zero when undisclosed and enters jointly with *MissingFunding*. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Among article-visible firms, public connections predict medical news evidence while their association with formal records is small and imprecise. Firms with a public connection have 2.16 times higher odds of medical news evidence; the linear probability estimate is 14.1 percentage points. These estimates

support Hypothesis 1. For formal records, the odds ratio is 1.12 and the linear estimate is 2.0 percentage points. The confidence intervals are wide, leaving Hypothesis 2 unresolved in these data.

Formal records rise with company age and funding. A one-standard-deviation increase in company age is associated with 2.41 times higher odds of a formal record, and the corresponding odds ratio for funding is 1.87. Both variables have smaller and imprecise associations with medical news evidence.

Article count is strongly associated with medical news evidence and weakly associated with formal records. This difference is expected because article count measures the amount of public text available for classifying the first outcome. The central comparison remains the public-connection coefficient: it is positive and precise for medical news evidence and close to one for formal records.

5.2. Robustness checks

The robustness checks ask whether the main results change when the sample expands, the public-connection measure changes, or the estimates are expressed in percentage points.

The first check estimates both outcome models on all 526 core firms. These models add the indicator for any retrieved article while keeping the remaining specification unchanged.

	Medical news evidence	Formal records
PublicConnection	2.27*** [1.32, 3.91]	1.20 [0.68, 2.11]
ArticleCount	2.98*** [1.88, 4.70]	0.84 [0.55, 1.28]
AnyArticle	8.64*** [2.94, 25.38]	3.06** [1.30, 7.21]
CompanyAge	1.02 [0.73, 1.41]	2.14*** [1.49, 3.05]
FundingAmount	1.06 [0.65, 1.72]	1.62** [1.06, 2.46]
MissingFunding	0.72 [0.31, 1.64]	1.49 [0.63, 3.54]
Observations	526	526
Positive outcomes	228	127
McFadden pseudo R^2	0.441	0.243
Funding-stage controls	Yes	Yes
Technology-family controls	Yes	Yes

Table 4. Full-core-sample logit estimates

*Notes: Cells report odds ratios with heteroskedasticity-robust 95 percent confidence intervals in brackets. AnyArticle indicates at least one retrieved public article. Other variable definitions and transformations follow Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

The full-core-sample estimates preserve the main separation. The public-connection odds ratio is 2.27 for medical news evidence and 1.20 for formal records. Company age and funding remain positive and precise in the formal-record model. The indicator for any retrieved article is positive in both outcome models, reflecting a broader difference between firms with and without searchable public text. The public-connection result persists when the analysis expands to the full core sample.

The second check replaces PublicConnection with alternative measures of network position. Table 5 reports separate full-core-sample models with the same controls.

Public-connection measure	Medical news evidence	Formal records
PublicConnection	2.27*** [1.32, 3.91]	1.20 [0.68, 2.11]
NeighborDegree	1.19 [0.91, 1.56]	1.07 [0.84, 1.38]
WeightedDegree	1.27 [0.95, 1.69]	1.09 [0.85, 1.40]
ConnectionPageRank	1.41** [1.06, 1.88]	1.07 [0.83, 1.38]
Observations	526	526
Funding-stage controls	Yes	Yes
Technology-family controls	Yes	Yes

Table 5. Alternative public-connection measures

Notes: Each cell reports a separate full-core-sample logit model in which the displayed measure replaces PublicConnection. PublicConnection is an indicator. The three continuous network measures enter as $\log(1 + x)$ standardized within the estimation sample. All models include ArticleCount, AnyArticle, CompanyAge, FundingAmount, MissingFunding, funding-stage controls, and technology-family controls. Cells report odds ratios with heteroskedasticity-robust 95 percent confidence intervals. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5 shows a consistent difference across outcomes. Neighbor degree, weighted degree, and PageRank are positive for medical news evidence, although only PageRank is estimated precisely. All three formal-record odds ratios remain close to one. The appendix also separates shared-person and shared-institution connections; the medical-news result is strongest for shared institutions (Table A3).

The third check uses linear probability models. Table A1 and Table A2 report a 14.1 percentage-point public-connection estimate for medical news evidence in the visible core sample and a 12.9-point estimate in the full core sample. The corresponding formal-record estimates are small and imprecise. Company age and funding remain positive and precise predictors of formal records.

Together, the checks show that public connections have outcome-specific predictive content. They are associated with whether clinical and regulatory activity enters public news, while formal records are

more closely associated with maturity and resources. This distinction suggests that public reporting and formal record production respond to different features of a startup and should be modeled as separate evidence processes.

6. Conclusion

This paper documents a consistent divergence across specifications in the public evidence surrounding Israeli digital health startups. Public connections predict medical news evidence among article-visible firms and in the full core sample. The formal-record estimate remains small and imprecise. Formal records rise with company age and funding.

The finding adds an evidence-system dimension to research on entrepreneurial signals. A public connection can facilitate activity, endorsement, and communication that generate medical news, while entry into clinical, regulatory, and scientific record systems requires accumulated organizational capacity. Treating these outcomes as separate processes yields a clearer account of what public connections reveal about young firms. Future work with event dates can test when connections form, when evidence appears, and how each predicts adoption, regulatory outcomes, financing, and survival.

Appendix A. Linear Probability Estimates

The linear probability estimates express the logit results in percentage points. Public connections predict medical news evidence in the visible and full core samples, while the formal-record estimates remain small. Company age and funding remain positive predictors of formal records.

	Medical news evidence	Formal records
PublicConnection	14.1*** (5.3)	2.0 (5.3)
ArticleCount	11.5*** (2.5)	-2.4 (2.8)
CompanyAge	0.8 (3.1)	15.3*** (3.6)
FundingAmount	3.6 (4.5)	11.3*** (4.3)
MissingFunding	-0.4 (9.1)	12.8* (7.5)
Observations	330	330
Funding-stage controls	Yes	Yes
Technology-family controls	Yes	Yes

Table A1. Linear probability estimates among article-visible firms

Notes: Cells report percentage-point coefficients with heteroskedasticity-robust standard errors in parentheses.

*Variable definitions and transformations follow Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

	Medical news evidence	Formal records
PublicConnection	12.9*** (4.3)	3.7 (4.4)
ArticleCount	16.1*** (3.1)	-2.8 (3.6)
AnyArticle	29.2*** (6.1)	13.3** (6.1)
CompanyAge	0.9 (2.1)	12.5*** (2.7)
FundingAmount	1.1 (3.2)	8.6** (3.6)
MissingFunding	-3.1 (4.7)	5.9 (4.7)
Observations	526	526
Funding-stage controls	Yes	Yes
Technology-family controls	Yes	Yes

Table A2. Full-core-sample linear probability estimates

Notes: Cells report percentage-point coefficients with heteroskedasticity-robust standard errors in parentheses.

AnyArticle indicates at least one retrieved public article. Other variable definitions and transformations follow

*Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

Appendix B. Public-Connection Diagnostics

Table A3 repeats the visible-core-sample logit models using alternative definitions of public connections. The broad public-connection indicator and the shared-institution indicator predict medical news evidence. Shared-person connections are imprecise. Formal-record estimates remain close to one across the alternative definitions.

Measure	Definition	Medical news evidence	Formal records
PublicConnection	Shared named person or institution	2.16*** [1.24, 3.77]	1.12 [0.60, 2.08]
SharedPersonConnection	Shared named person	1.36 [0.59, 3.15]	1.03 [0.41, 2.59]
SharedInstitutionConnection	Shared named institution	1.91** [1.06, 3.44]	1.12 [0.60, 2.09]
PublicCitableConnection	Shared public-citable named person or institution	1.93** [1.10, 3.41]	1.15 [0.61, 2.17]
Observations		330	330
Funding-stage controls		Yes	Yes
Technology-family controls		Yes	Yes

Table A3. Alternative definitions of public connections

*Notes: Each row reports a separate article-visible-sample logit model in which the displayed definition replaces PublicConnection. All models include ArticleCount, CompanyAge, FundingAmount, MissingFunding, funding-stage controls, and technology-family controls. Cells report odds ratios with heteroskedasticity-robust 95 percent confidence intervals. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

Appendix C. Extended Public-News Outcomes

The extended models compare medical news evidence with financing-news and partnership-news evidence in the same full core sample. These outcomes broaden the descriptive comparison while leaving the paper's two central outcomes unchanged.

	Financing news evidence	Partnership news evidence	Medical news evidence	Formal records
PublicConnection	1.30 [0.74, 2.27]	0.88 [0.44, 1.77]	2.27*** [1.32, 3.91]	1.20 [0.68, 2.11]
ArticleCount	2.95*** [1.78, 4.88]	5.22*** [2.54, 10.71]	2.98*** [1.88, 4.70]	0.84 [0.55, 1.28]
AnyArticle	9.78*** [3.56, 26.86]	13.68*** [4.42, 42.30]	8.64*** [2.94, 25.38]	3.06** [1.30, 7.21]
CompanyAge	0.81 [0.58, 1.13]	0.93 [0.63, 1.37]	1.02 [0.73, 1.41]	2.14*** [1.49, 3.05]
FundingAmount	1.40 [0.87, 2.25]	0.83 [0.42, 1.64]	1.06 [0.65, 1.72]	1.62** [1.06, 2.46]
MissingFunding	0.66 [0.32, 1.36]	0.56 [0.21, 1.54]	0.72 [0.31, 1.64]	1.49 [0.63, 3.54]
Observations	526	526	526	526
Funding-stage controls	Yes	Yes	Yes	Yes
Technology-family controls	Yes	Yes	Yes	Yes

Table A4. Extended full-core-sample logit estimates

*Notes: Cells report odds ratios with heteroskedasticity-robust 95 percent confidence intervals in brackets. All four models use the same full core sample. Variable definitions and transformations follow Table 4. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

Full-core-sample logit estimates

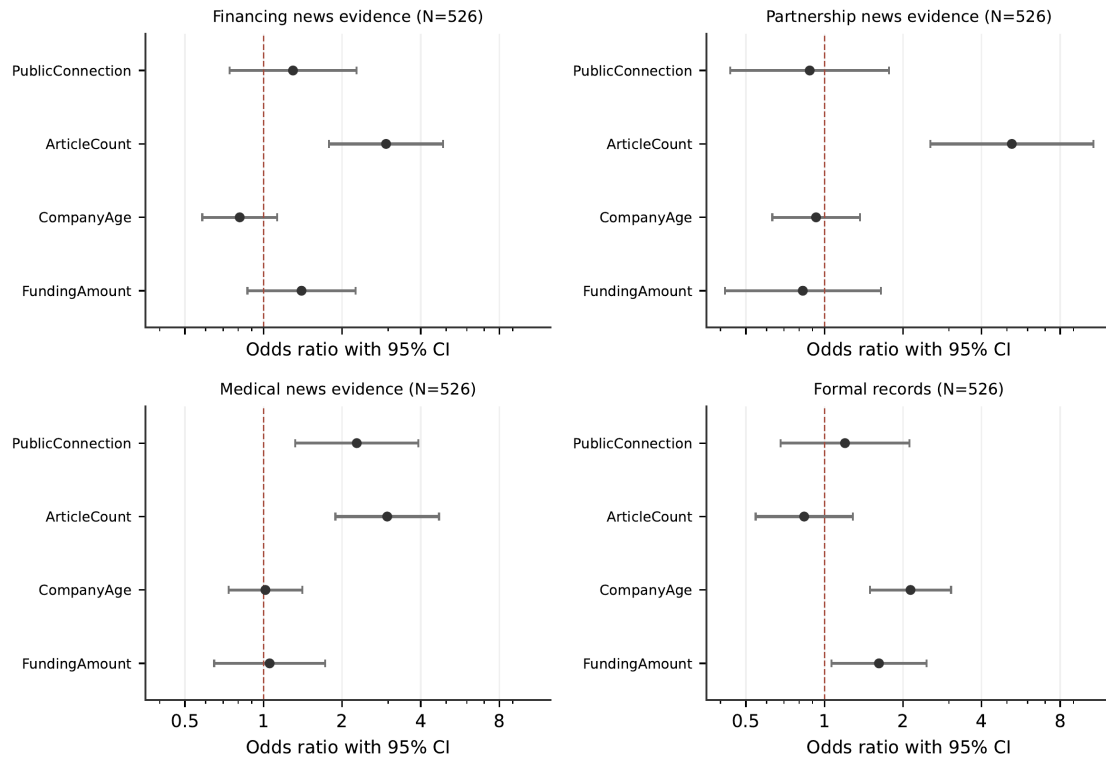


Figure A1. Full-core-sample logit estimates for four evidence outcomes

Figure A1 labels all four models and their common sample size. Each point is an odds ratio, each horizontal bar is a heteroskedasticity-robust 95 percent confidence interval, and the dashed vertical line marks an odds ratio of one. Article count is strongly associated with all three news outcomes. Company age and funding are most strongly associated with formal records.

Statements and Declarations

Data Availability and Replication

The company-level dataset combines public company profiles, public news, public connections, and formal records from clinical, regulatory, and research systems. Source-access and redistribution constraints limit public release of the underlying dataset. The dataset is available from the authors for research review subject to those constraints. The replication materials contain the analysis code, derived tables and figures, model outputs, and documentation used in this manuscript.

Acknowledgments

We gratefully acknowledge support from the Soft Science Research Program of Sichuan Provincial Department of Science and Technology (2025NSFSCR0089).

Footnotes

¹ <https://finder.startupnationcentral.org/>

References

1. [△]Akerlof GA (1970). "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *Q J Econ.* **84**(3):488–500.
2. [△][△]Conti A, Thursby J, Thursby MC (2013). "Patents as Signals for Startup Financing." *J Ind Econ.* **61**(3):592–622.
3. [△][△][△]Hoenig D, Henkel J (2015). "Quality Signals? The Role of Patents, Alliances, and Team Experience in Venture Capital Financing." *Res Policy.* **44**(5):1049–1064.
4. [△][△][△]Hsu DH (2004). "What Do Entrepreneurs Pay for Venture Capital Affiliation?" *J Finance.* **59**(4):1805–1844.
5. [△][△]Kaplan SN, Sensoy BA, Strömberg P (2009). "Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies." *J Finance.* **64**(1):75–115.
6. [△]Spence M (1973). "Job Market Signaling." *Q J Econ.* **87**(3):355–374.
7. [△][△][△]Stuart TE, Hoang H, Hybels RC (1999). "Interorganizational Endorsements and the Performance of Entrepreneurial Ventures." *Adm Sci Q.* **44**(2):315–349.
8. [△]Startup Nation Central (2026). "Startup Nation Finder." Startup Nation Central. <https://finder.startupnationcentral.org/>.
9. [△]ClinicalTrials.gov (2026). "ClinicalTrials.gov Data API." ClinicalTrials.gov. <https://clinicaltrials.gov/data-api>.
10. [△]U.S. Food and Drug Administration (2026). "OpenFDA Device 510(k) Clearances API." U.S. Food and Drug Administration. <https://open.fda.gov/apis/device/510k/>.
11. [△]OpenAlex (2026). "OpenAlex API Documentation." OpenAlex. <https://developers.openalex.org/>.

Supplementary data: available at <https://doi.org/10.32388/R6XYNH>

Declarations

Funding: Soft Science Research Program of Sichuan Provincial Department of Science and Technology (2025NSFSCR0089)

Potential competing interests: No potential competing interests to declare.