

THE IMPACT OF DATA BREACH COSTS ON MACROECONOMIC INDICATORS: AN EMPIRICAL PANEL DATA ANALYSIS

Astghik PARSYAN

PhD in Economics, Assistant, at YSU Faculty of Economics and Management

Key words: data breach, data breach costs, cybersecurity, employment, GDP, macroeconomic impact, panel data

Introduction

The deepening digitalization of the economy and the rapid development of information technologies have transformed data into a key resource of economic activity. As the volume of data continues to grow, so too do the risks of unauthorized access, cyberattacks, and data breaches. Consequently, cybersecurity is gradually moving beyond the scope of a purely technical issue to acquire economic significance, since the consequences of cyber incidents can affect organizations' financial performance, investment activity, productivity, and, more broadly, economic growth. The economic scale of the data breach problem can also be assessed through its associated costs. According to IBM's 2025 "Cost of a Data Breach" report [IBM, 2025], the global average cost of a single data breach amounted to USD 4.44 million, compared to USD 4.88 million in the previous year. At the same time, research conducted by the World Bank [Cobos, 2024] indicates that the level of cybersecurity may be associated not only with the performance of individual organizations but also with the economic performance of countries; in particular, a substantial reduction in the level of cyber incidents in developing countries may, in the long run, contribute to growth in per capita GDP. The economic consequences of data breaches can manifest both through direct financial losses and through indirect costs, including business disruption, additional expenses for recovery and security enhancement, changes in investment decisions, and potential productivity losses. The magnitude of these effects may vary depending on countries' level of economic development and digitalization, their institutional environment, and their cybersecurity capabilities. Consequently, the quantitative assessment of the relationship between data breach costs and macroeconomic indicators is important for comprehensively evaluating the economic significance of the problem.

The purpose of this study is to employ econometric methods to assess the impact of data breach costs on countries' macroeconomic indicators. The object of the study is the costs associated with data breaches at the country level, while the subject is their impact on countries' macroeconomic indicators. Within the scope of the study, panel data are used to assess the impact of data breach costs on both gross domestic product and the level of employment. The study is based on the hypothesis that an increase in data breach costs may have a negative effect on both GDP and the level of employment. For the empirical analysis, data from international sources on data breaches are combined with countries' macroeconomic statistical data.

Methodology

The study was conducted using annual data for the period 2017–2024 for 12 countries: the United States, Canada, the United Kingdom, France, Germany, Italy, Brazil, India, Japan, South Korea, Australia, and South Africa. The overall sample comprises 96 country-year observations. Information on data breach costs was obtained from IBM's annual Cost of a Data Breach Report, while data on macroeconomic indicators were obtained from the World Bank database.

The main explanatory variable of the study is data breach costs (DB), expressed in millions of USD. To assess their impact, two panel econometric models were constructed. In Model 1, the dependent variable is the number of employed persons (EMP), with investment (INV), household consumption expenditure (HHCE), and the volume of exports of goods and services (EXP) included as additional explanatory variables. In Model 2, the dependent variable is nominal gross domestic product (GDP), with the volume of exports (EXP) included as an additional explanatory variable. Indicators expressed in monetary terms were measured in millions of USD. The use of panel data makes it possible to simultaneously account for differences across the countries under study, for variation over time.

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \dots + \beta_k X_{kit} + \varepsilon_{it}$$

In the panel data structure, (i) denotes the country under study, while (t) denotes the corresponding year of observation. The index (it) for each variable therefore indicates the value of that variable for a specific country in a specific year [Gujarati & Porter, 2009, 592–606].

Prior to estimating the panel models, the stationarity of the time series under study was tested. For this purpose, the Augmented Dickey-Fuller–Fisher (ADF-Fisher) unit root test was applied. The null hypothesis of the test assumes the presence of a unit root; when ($p < 0.05$), the null hypothesis is rejected, and the series is considered stationary. For non-stationary series, the first-difference operator was applied to obtain a stationary series.

For each model, Pooled Ordinary Least Squares (Pooled OLS), Fixed Effects (FE), and Random Effects (RE) panel models were first estimated. The Fixed Effects approach makes it possible to account for country-specific, time-invariant unobserved factors. In the Random Effects approach, the country-specific individual effect is treated as a random component and is assumed to be independent of the model's explanatory variables. The choice between the Fixed Effects and Random Effects models was made using the Hausman test. The null hypothesis of the test assumes that the country-specific individual effect is independent of the explanatory variables, in which case the Random Effects estimator is considered an efficient and consistent estimation approach. When ($p > 0.05$), the null hypothesis is not rejected, and the Random Effects model is preferred. Possible cross-sectional dependence among the models' residuals was assessed using the Residual Cross-Section Dependence Test. For this purpose, the Breusch–Pagan LM, Pesaran scaled

LM, and Pesaran CD tests were applied. Based on the results of these tests, it was assessed whether statistically significant interdependence exists among the residuals of the countries under study. The final models were estimated using the Panel EGLS method with a cross-section random effects approach. Panel-Corrected Standard Errors (PCSE) were applied for standard error estimation and for constructing the covariance matrix. The statistical significance of the individual model coefficients was assessed using t-statistics and the corresponding p-values, while the overall significance of the models was assessed using the F-statistic and its p-value. The explanatory power of the models was evaluated using the R^2 and Adjusted R^2 statistics. The final model specifications, the results of the diagnostic tests, and the econometric estimates are presented below.

Literature review

In the literature the economic consequences of data breaches and cybercrime have been examined at both the organizational and macroeconomic levels. Prior studies have primarily assessed their impact on productivity, labor, consumption, return on capital, and other indicators of economic activity. Makridis and Dean (2018), in their article "Measuring the Economic Effects of Data Breaches on Firm Outcomes: Challenges and Opportunities," [Makridis & Dean, 2018, 59–83] examined the impact of data breaches on organizations' economic outcomes. The authors constructed a database combining publicly recorded information on data breaches with firms' financial performance. In the first stage of the study, they identified substantial differences in the prevalence and scale of data breaches across industries, also between public and private companies. In the second stage, by estimating firm-level productivity, the authors found, within certain model specifications, that a 10% increase in data breaches is associated with an approximately 0.2% decline in firm productivity. At the same time, the authors emphasize that this result is sensitive to the choice of database and model specification.

Tetteh and Otioma (2025), in their article "Cyberattack, Cyber Risk Mitigation Capabilities, and Firm Productivity in Kenya," [Tetteh & Otioma, 2025, 1493–1514] examined the impact of cybersecurity breaches on labor productivity. The study used data from Kenya's 2016 Enterprise ICT Survey. The authors applied Principal Component Analysis to assess cyber risk mitigation capabilities, while employing Ordinary Least Squares regression in the main empirical estimations, while also accounting for potential endogeneity concerns. The results showed that cybersecurity breaches have a negative effect on firms' labor productivity. At the same time, it was found that employee training and the development of digital skills can mitigate this negative effect, particularly among small and medium-sized enterprises. The authors also showed that information technology policy and technical security measures can contribute to increased productivity, but are not, on their own, sufficient to fully offset the adverse effects of cyberattacks.

He (2026), in the study "Unveiling the Socio-Economic Toll of Cybercrime: Evidence from South Korea," [He, 2026] examined the economic impact of cybercrime at the macro-economic level, focusing on the South Korean economy. The author applied Bayesian estimation and impulse response functions to assess the impact of cybercrime on key economic variables. The results showed that cybercrime shocks have a negative effect on household consumption, firm productivity, and return on capital. At the same time, the study shows that public investment in cybersecurity can mitigate the short-term negative effects of cyber-crime and, in the long run, contribute to greater economic resilience.

The authors of "The Economic Impact of Cyberattacks" (2026) [Uguet & Jung, 2026] also examined the impact of cyberattacks on economic activity and GDP. Within the framework of this study, the authors developed a nonlinear theoretical model that represents the economic impact of cyberattacks through productivity and GDP. For the empirical estimation, the authors used data from a range of countries and applied nonlinear least squares and the Generalized Method of Moments (GMM) in order to account for potential endogeneity issues. The results obtained indicate that cyberattacks can reduce workers' economic contribution, as time and resources are diverted toward system recovery or the replacement of digital processes with manual labor. The authors also found some evidence of a decline in labor productivity and estimated the economic loss from cyberattacks for the countries studied to average around 0.32% of GDP.

The theoretical and empirical foundations for assessing the economic costs of cyber incidents are systematically presented in Estefania Vergara Cobos and Selcen Cakir's (2024) study "A Review of the Economic Costs of Cyber Incidents," published by the World Bank [Vergara Cobos & Cakir, 2024]. The authors synthesize the existing empirical literature on the economic consequences of cyber incidents and distinguish between their direct and indirect costs. According to the study, assessing the economic impact of cyber incidents is complicated by wide variation in estimates and limited data availability. The study emphasizes that economic damage is not limited to immediate financial losses; it may include costs from business disruption, productivity losses, and other secondary effects.

Scientific novelty

The scientific novelty of this study lies in the quantitative assessment of the relationship between data breach costs and a country's macroeconomic indicators, based on panel data from 12 countries for the period 2017–2024.

An important component of this novelty is the assessment of the impact of data breach costs along two dimensions—changes in the number of employed persons and changes in nominal GDP—as well as the identification of the temporal pattern of this impact. In particular, the findings of the study show that the economic consequences of data breach costs may also manifest with a certain time lag.

Analysis

Within the framework of Model 1, the impact of data breach costs on changes in the number of employed persons was assessed, controlling for changes in investment, household consumption expenditure, and export volumes. As a result of the estimation—taking into account the first-difference transformation applied to non-stationary series and the one-period lag of the export variable—the final sample comprised 72 country-year observations for the period 2019–2024. The choice between the Fixed Effects and Random Effects approaches for model estimation was made using the Hausman test. The test results showed a Chi-square statistic of 0.0000, with a corresponding p-value of 1.0000. Since $p > 0.05$, the null hypothesis of no systematic difference between the Fixed Effects and Random Effects estimates is not rejected. Accordingly, the Random Effects approach is for Model 1.

Table 1. Model 1: Random Effects Model

Dependent Variable: D(LEMP_)				
Method: Panel EGLS (Cross-section random effects)				
Sample (adjusted): 2019 2024				
Periods included: 6				
Cross-sections included: 12				
Total panel (balanced) observations: 72				
Swamy and Arora estimator of component variances				
Cross-section weights (PCSE) standard errors & covariance (d.f.corrected)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.0078	0.0037	2.0955	0.0399
D(DB_)	-0.0161	0.0058	-2.7498	0.0077
D(INV_)	1.49E-07	3.68E-08	4.0615	0.0001
D(LHHCE_)	0.0689	0.0294	2.3403	0.0223
D(LEXP_(-1))	0.1023	0.0267	3.8292	0.0003
Effects Specification				
			S.D.	Rho
Cross-section random			0.0073	0.0924
Idiosyncratic random			0.0230	0.9076
Weighted Statistics				
R-squared	0.4277	Mean dependent var		0.0107
Adjusted R-squared	0.3936	S.D. dependent var		0.0298
S.E. of regression	0.0232	Sum squared resid		0.0361
F-statistic	12.5227	Durbin-Watson stat		1.9671
Prob(F-statistic)	0.0000			
Unweighted Statistics				
R-squared	0.4135	Mean dependent var		0.0135
Sum squared resid	0.0400	Durbin-Watson stat		1.7751

The estimated model equation is:

$$D(LEMP) = 0.0078 - 0.0161D(DB) + 1.49E-07D(INV) + 0.0689D(LHHCE) + 0.1023D(LEXP(-1))$$

The coefficient on data breach costs (DB) is -0.0161 and is statistically significant ($p = 0.0077$). The results indicate that a USD 1 million increase in data breach costs is associated with an approximately 1.6% decrease in the number of employed persons, holding other conditions constant. This negative and statistically significant effect confirms the study's first hypothesis. The coefficient on investment (INV) is positive and statistically significant, at 1.49×10^{-7} ($p = 0.0001$). A USD 1 million increase in investment is associated with an approximately 0.0015% increase in the number of employed persons. Despite the positive direction of the effect, its magnitude is small.

The coefficient on household consumption expenditure (HHCE) is 0.0689 and is statistically significant ($p = 0.0223$). The results show that a 1% increase in consumption expenditure is associated with an approximately 0.069% increase in the number of employed persons. The exports coefficient (EXP) is 0.1023 and is statistically significant ($p = 0.0003$). This result indicates that a 1% increase in exports in the previous period is associated with an approximately 0.102% increase in the number of employed persons. Thus, export expansion is also positively associated with changes in the number of employed persons.

The overall statistical indicators confirm the significance of the model. The R^2 is 0.4277, and the Adjusted R^2 is 0.3936, meaning that the variables included in the model explain approximately 42.8%, and, on an adjusted basis, 39.4%, of the variation in the number of employed persons. The overall significance of the model is confirmed by the F-statistic of 12.5227, with a p-value of 0.0000.

Overall, the results of Model 1 show that an increase in data breach costs has a negative and statistically significant effect on the number of employed persons, whereas increases in investment, household consumption expenditure, and exports in the previous period are associated with a positive change in the number of employed persons.

Within the framework of Model 2, the impact of data breach costs on changes in nominal gross domestic product was assessed, with export volume and the previous period's change in GDP included as control variables. As a result of the estimation—taking into account the first-difference transformation applied to non-stationary series and the inclusion of lagged variables for data breach costs and GDP—the final sample comprised 48 country-year observations. The choice between the Fixed Effects and Random Effects approaches for model estimation was made using the Hausman test. The test results yielded a Chi-square statistic of 5.2785, with a corresponding p-value of 0.2599. Since $p > 0.05$, the null hypothesis is not rejected, and no statistically significant systematic difference between the Fixed Effects and Random Effects estimates is found. Accordingly, the Random Effects approach was selected for the estimation of Model 2.

Possible cross-country dependence among this model's residuals was tested using the Residual Cross-Section Dependence Test. The Breusch–Pagan LM test statistic was

82.1706 ($p = 0.0864$), the Pesaran scaled LM test statistic was 1.4074 ($p = 0.1593$), and the Pesaran CD test statistic was 0.8015 ($p = 0.4228$). In all cases, the p-values exceed 0.05, and thus the presence of statistically significant cross-sectional dependence among the residuals across countries is not confirmed.

The final estimation of Model 2 was carried out using the Panel EGLS method with a cross-section random effects approach. The dependent variable is the change in nominal GDP, and the model includes the current change in export volume, the second- and third-period lagged changes in data breach costs, and the second-lagged change in GDP.

Table 2. Model 2: Random Effects Model

Dependent Variable: D(LGDP_)				
Method: Panel EGLS (Cross-section random effects)				
Periods included: 4				
Cross-sections included: 12				
Total panel (balanced) observations: 48				
Swamy and Arora estimator of component variances				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011	0.0112	0.978	0.3335
D(LEXP_)	0.5742	0.0795	7.223	0.0000
D(LDB_(-2))	-0.2206	0.0631	-3.497	0.0011
D(LDB_(-3))	-0.2518	0.1001	-2.515	0.0157
D(LGDP_(-2))	0.4595	0.1247	3.6851	0.0006
Effects Specification			S.D.	Rho
Cross-section random			0	0
Idiosyncratic random			0.0498	1
Weighted Statistics				
R-squared	0.6408	Mean dependent var		0.0534
Adjusted R-squared	0.6074	S.D. dependent var		0.0773
S.E. of regression	0.0484	Sum squared resid		0.1008
F-statistic	19.1756	Durbin-Watson stat		2.3331
Prob(F-statistic)	0.0000			
Unweighted Statistics				
R-squared	0.6408	Mean dependent var		0.0534
Sum squared resid	0.1008	Durbin-Watson stat		2.3331

The estimated model equation is: $D(LGDP) = 0.011 + 0.5742D(LEXP) - 0.2206D(LDB(-2)) - 0.2518D(LDB(-3)) + 0.4595D(LGDP(-2))$

The coefficient on exports (EXP) is 0.5742 and is statistically significant ($p = 0.0000$). The results indicate that a 1% increase in exports is associated with an approximately 0.574% increase in nominal GDP, holding other conditions constant. Thus, export expansion has a

positive and statistically significant relationship with changes in GDP. The coefficient on the second-period lagged change in data breach costs (DB) is -0.2206 and is statistically significant ($p = 0.0011$). This means that a 1% increase in data breach costs is associated with an approximately 0.221% decrease in nominal GDP two periods later. This result points to a certain time lag in the economic impact of data breaches and indicates that their negative effect on GDP may not manifest immediately. The coefficient on the third-period lagged change in data breach costs is likewise negative and statistically significant, at -0.2518 ($p = 0.0157$). Accordingly, a 1% increase in data breach costs is associated with an approximately 0.252% decrease in nominal GDP three periods later. The negative and significant coefficients for both the second- and third-period lags indicate that the economic consequences of data breaches may persist over time.

The coefficient on the second-period lagged change in GDP is 0.4595 and is statistically significant ($p = 0.0006$). This result indicates that a 1% increase in GDP in the previous period is associated with an approximately 0.460% increase in current GDP, reflecting a certain continuity in economic activity over time. The overall statistical indicators confirm the model's high overall significance and adequate explanatory power. The R^2 is 0.6408 , and the Adjusted R^2 is 0.6074 , meaning that the variables included in the model explain approximately 64.1%, and, on an adjusted basis, 60.7%, of the variation in nominal GDP. The overall statistical significance of the model is confirmed by the F-statistic of 19.1756 , with a p-value of 0.0000 . Overall, the results of Model 2 show that an increase in data breach costs has a negative and statistically significant relationship with nominal GDP, which manifests after two and three periods. Also, increases in exports and in economic activity in the previous period are associated with a positive change in nominal GDP.

Conclusion

The purpose of this study was to assess the impact of data breach costs on key indicators of economic activity, based on panel data from 12 countries for the period 2017–2024. Two models were estimated as part of this work, through which the impact of data breach costs was examined with respect to changes in the number of employed persons and in nominal GDP, respectively. The findings of the study show that an increase in data breach costs is statistically significantly associated with a decrease in the number of employed persons. At the same time, the results of Model 2 show that an increase in data breach costs has a negative relationship with changes in nominal GDP, with this effect remaining statistically significant in subsequent periods as well. This finding suggests that the economic consequences of data breaches may not be confined to the immediate period of the incident and may instead manifest with a certain time lag. These results support the study's central conclusion: data breaches should be viewed not merely as an organizational or cybersecurity issue, but as a factor of economic risk. Their consequences can extend to the labor market and to overall economic activity, generating an impact that is not fully captured by immediate financial losses alone. From this perspective, cybersecurity and data protection

measures also acquire economic significance. The prevention of data breaches and the rapid mitigation of their consequences can be viewed not only as a matter of information security but also as a means of managing economic risk. Overall, the findings of the study confirm the empirical validity of the main hypothesis put forward and indicate that the economic impact of data breach costs can be both multidimensional and persistent over time. This underscores the need to consider cybersecurity issues within the broader context of economic policy and stability.

References

1. IBM. (2025). *Cost of a Data Breach Report 2025*. IBM. <https://www.ibm.com/reports/data-breach>
2. World Bank. (2025). *Cybersecurity Economics for Emerging Markets*. World Bank.
3. Makridis, C., & Dean, B. (2018). Measuring the economic effects of data breaches on firm outcomes: Challenges and opportunities. *JESM*, 43(1–2), 59–83. <https://doi.org/10.3233/JEM-180450>
4. Tetteh, G. K., & Otioma, C. (2025). Cyberattack, cyber risk mitigation capabilities, and firm productivity in Kenya. *Small Business Economics*, 64(3), 1493–1514. <https://doi.org/10.1007/s11187-024-00946-8>
5. He, Y. (2026). Unveiling the socio-economic toll of cybercrime: Evidence from South Korea. *Humanities and Social Sciences Communications*, 13, Art. 242. <https://doi.org/10.1057/s41599-026-06521-8>
6. The Economic Impact of Cyberattacks. (2026). *Journal of Digital Economy*. <https://doi.org/10.1016/j.jdec.2026.07.001>
7. Vergara Cobos, E., Cakir, S. (2024). *A Review of the Economic Costs of Cyber Incidents*. Washington, DC: WB.
8. Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics* (5th ed.). Boston: McGraw-Hill/Irwin, pp. 591–606.

Astghik PARSYAN

The impact of data breach costs on macroeconomic indicators: an empirical panel data analysis

Key words: data breach, data breach costs, cybersecurity, employment, GDP, macroeconomic impact, panel data

This paper examines the impact of data breach costs on employment and nominal gross domestic product. The study is based on panel data from 12 countries covering the period 2017–2024. Panel data methods were employed for econometric estimation, and the final model selection was performed using the Hausman test. The estimation results indicate that the Random Effects approach is preferable for both models. The findings of the study show that an increase in data breach costs has a statistically significant negative effect on employment. At the same time, it was found that the negative impact of data breach costs on GDP manifests with a time lag. These results suggest that data breaches may generate not only immediate financial losses but also adverse effects on the real sector of the economy and on key indicators of macro-economic development. The scientific novelty lies in the quantitative assessment of the macroeconomic consequences of data breach costs, examined through their impact on employment and GDP within a panel data framework.