International Journal of Scientific Research and Modern Education (IJSRME)
Impact Factor: 7.137, ISSN (Online): 2455 - 5630
(www.rdmodernresearch.com) Volume 7, Issue 2, 2022

TRANSACTION CHARACTERISTICS AND FINANCIAL FRAUD DETECTION IN SIMULATED BANK PAYMENT SYSTEMS USING PUBLIC BANKSIM DATA

Divya Vishwanath* & Paul Johnson Asamoah**
* Janardhan Rai Nagar Rajasthan Vidyapeeth, Udaipur, Rajasthan, India
** Power Life University College, Accra, Ghana, West Africa

Cite This Article: Divya Vishwanath & Paul Johnson Asamoah, "Transaction Characteristics and Financial Fraud Detection in Simulated Bank Payment Systems Using Public Banksim Data", International Journal of Scientific Research and Modern Education, Volume 7, Issue 2, Page Number 84-92, 2022.
Copy Right: © IJSRME, 2022 (All Rights Reserved). This is an Open Access Article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract:

This study examines how transaction characteristics relate to fraud labels in simulated bank payments, addressing the need to interpret fraud signals within their payment context. The analysis uses 17,200 BankSim transactions across 180 simulation steps and estimates a weighted linear probability model with customer clustered standard errors. A 100 unit increase in transaction amount is associated with a 4.27 percentage point increase in fitted fraud probability after adjustment for payment category and customer age code. Fraud status also differs across payment categories, while the age codes show no detectable joint association. The positive amount association remains in an exploratory analysis restricted to categories containing both fraud and nonfraud records. These findings refine transaction anomaly reasoning by indicating that payment value and category provide distinct information within the simulated data. The associations do not establish causal mechanisms or operational detection performance. Because some fitted values fall outside the probability range, the estimates should not be used as individual risk probabilities. Financial institutions can use the findings to guide the evaluation of transaction characteristics against representative operational data before adopting fraud alerts.

Key Words: Banksim; Fraud Detection; Payment Transactions; Sampling Weights; Transaction Characteristics

1. Introduction:

Financial payment systems must distinguish fraudulent activity from legitimate transactions without treating every unusual payment as suspicious. This distinction affects how institutions allocate investigation resources and how customers experience payment controls. The challenge is not only to identify transactions that differ from ordinary activity, but also to determine which characteristics remain informative when considered together. Dal Pozzolo et al. 2014 explain that fraud detection operates under changing transaction patterns and an imbalance between legitimate and fraudulent payments. These conditions make the interpretation of transaction evidence central to the design of detection systems.

Transaction amount offers one potential source of information because fraudulent and legitimate payments may follow different value distributions. A fraudulent actor may seek a high return from an unauthorized payment, while a smaller payment may be selected to attract less attention. Legitimate transactions can also have high or low values. The expected relationship between amount and fraud status is therefore an empirical question rather than a rule that applies to every payment. Correa Bahnsen et al. 2016 show that the representation of transaction characteristics matters when constructing fraud detection models. Their work also indicates that the usefulness of a characteristic depends on how it is measured and combined with other information.

Payment category provides a second dimension of transaction context. The value and frequency of legitimate payments can differ across categories, and simulated fraudulent activity may also be distributed unevenly across them. A transaction amount that is common in one category may therefore have a different meaning in another. Research by Jha et al. 2012 uses aggregated transaction information to examine fraud detection, while Van Vlasselaer et al. 2015 combine transaction characteristics with information on customer and merchant relationships. Together, these approaches demonstrate that transaction context can be represented at different levels. They do not establish that any category difference observed in a simulated dataset will transfer to an operating bank.

Customer age code provides a third observable characteristic, although its relationship with fraud status is less direct. Differences across age categories may reflect variation in simulated purchasing patterns or in the transactions assigned to those categories. Such differences need not persist after transaction amount and payment category are considered. Nor would an association establish that age itself causes fraud exposure. These distinctions are important for interpreting demographic information: a recorded customer characteristic may describe the composition of transactions without identifying the behavior that produced a fraud label.

The outcome examined here is the fraud label assigned to each simulated payment. It identifies whether a transaction is recorded as fraudulent or nonfraudulent in the source dataset. This measure supports an analysis

of associations with transaction characteristics, but it is not a measure of detection speed, investigation quality, financial loss, or the performance of an operational alert system. Stojanović et al. 2021 analyze BankSim among several fraud datasets and document its transaction fields and simulated fraud labels. Their work establishes the dataset as an available setting for fraud analysis while also showing that results depend on the data and analytical methods used.

The closest research has demonstrated the use of transaction features, aggregated spending information, and customer and merchant relationships in fraud detection. The reviewed evidence does not, however, establish how transaction amount, payment category, and customer age code are jointly associated with the fraud label in the particular weighted sample analyzed here. The contribution is a bounded assessment of those associations within BankSim, including the effect of the sampling design on interpretation. It does not depend on a claim that the dataset, its variables, or the general fraud detection problem are new.

This study aims to examine the associations between transaction characteristics and recorded fraud status in simulated bank payments. The research question asks how transaction amount, payment category, and customer age code relate to the likelihood that a payment is labeled fraudulent. The analysis tests whether higher transaction amounts are associated with fraud status and examines whether fraud status differs across payment categories and customer age codes.

The empirical analysis uses 17,200 simulated transactions from BankSim, a dataset calibrated using aggregate payment information supplied by a bank in Spain. The sample retains all 7,200 transactions labeled fraudulent and a random sample of 10,000 nonfraudulent transactions. A weighted linear probability model estimates the relationships, with standard errors clustered by customer. The analysis finds a positive adjusted association for transaction amount and joint differences across payment categories, but no detectable joint association for age code. The model also produces fitted values outside the probability range for some transactions, limiting its use for individual prediction.

The findings offer a basis for examining transaction characteristics jointly and for assessing how sampling changes the interpretation of fraud patterns. Their transfer to other payment systems remains an empirical question because the data are simulated and do not measure live detection performance. The remainder of the article presents the theoretical framework and research questions in Section 2, the data in Section 3, and the methods in Section 4. Section 5 reports the findings, and Section 6 discusses their significance.

2. Theory, Conceptual Framework and Research Questions Development:

Financial fraud detection depends on distinguishing fraudulent payments from legitimate transactions within a system that connects customers and merchants. Each payment produces observable information, including its amount, payment category, and customer age code. These characteristics may help distinguish simulated fraud from legitimate activity when the two groups follow different patterns. However, legitimate and fraudulent transactions can also share the same characteristics, limiting the value of any single indicator. Dal Pozzolo et al. 2018 show that fraud detection must account for differences between fraudulent and legitimate transactions and for changes in transaction patterns over time.

This analysis draws on the logic of transaction anomaly detection. The theoretical expectation is that fraudulent payments may differ from legitimate payments because their objectives and selection of transactions differ. The outcome is the recorded fraud status of each simulated payment. Transaction amount, payment category, and customer age code are observed characteristics used to examine that outcome. Stojanović et al. 2021 demonstrate the use of these characteristics in BankSim fraud detection. Their analysis supports their inclusion as measurable predictors but does not establish a causal effect for any individual characteristic.

Transaction amount represents the monetary value of a payment. A fraudulent actor may seek a higher return from an unauthorized transaction, creating a theoretical expectation that larger payments are more likely to carry a fraud label. Conversely, fraudulent payments may be small when an actor seeks to avoid attention. Legitimate high value payments may also occur, so amount alone cannot establish fraud. Research by Jha and Westland 2013 supports examining patterns in fraudulent card transactions, while leaving the direction of any association dependent on the transactions studied.

H1: Higher transaction amounts are associated with a greater likelihood of a simulated payment being labeled fraudulent.

Payment category identifies the type of purchase associated with a transaction. Categories may differ in typical payment values and in the opportunities, they offer for fraudulent purchases. These differences can produce variation in fraud labels across categories. The direction for a particular category depends on the reference category and the simulated payment patterns. Seera et al. 2021 examine transaction information, including merchant category, as part of payment card fraud detection. Their work supports assessing category as a predictor but does not establish that every category has the same relationship with fraud.

RQ1: Does the likelihood of a simulated payment being labeled fraudulent differ across payment categories?

Customer age code identifies the age category assigned to the customer associated with a payment. Age categories may differ in their simulated purchasing patterns and exposure to payment categories, which can produce differences in recorded fraud status. Fraudulent activity may also occur across age categories, and the available codes do not establish the behavior or vulnerability of individual customers. Randhawa et al. 2018 show that fraud detection performance depends on the data and methods used, supporting an empirical assessment rather than a predetermined age gradient.

RQ2: Does the likelihood of a simulated payment being labeled fraudulent differ across customer age codes?

The conceptual framework treats transaction amount, payment category, and customer age code as observed predictors of recorded fraud status. The proposed relationships are associations. The analysis can find support, no support, or inconclusive evidence for H1 and can identify differences or no differences for the two research questions. The fraud records are simulated, and the sample contains a larger proportion of fraudulent payments than the source dataset. Sampling weights are therefore needed when estimating fraud rates for the source dataset. Findings describe the simulated transactions and do not establish causal relationships in actual banking systems.

3. Data:

This analysis uses simulated bank payments to examine whether transaction characteristics are associated with recorded fraud status. The data section documents the source, sample construction, observed variables, and analytical requirements.

3.1 Data Sources, Population, and Sample Construction:

The analysis uses the BankSim transaction dataset, generated from a simulation calibrated with aggregate payment information supplied by a bank in Spain. Stojanović et al. 2021 describe BankSim as a synthetic dataset containing legitimate and fraudulent payments for fraud detection research.

The unit of analysis is a simulated payment transaction. The accessible frame contains 594,643 transactions across 180 simulation steps, numbered 0 to 179. The data represent repeated transactions associated with customers and merchants in a simulated banking payment system. The simulation steps do not identify exact calendar dates. The target population for inference is the set of transactions generated in this BankSim dataset.

The analytical sample retains all 7,200 transactions labeled fraudulent and selects 10,000 of the 587,443 transactions labeled nonfraudulent through random sampling. The resulting sample contains 17,200 transactions associated with 3,927 customer identifiers and 49 merchant identifiers. The source file was accessed through a public repository mirror. Its release version and applicable licence require confirmation from the dataset provider. The extraction date is recorded in the accompanying workbook.

Sample Construction Stage	Transactions	Selection Rule
Accessible BankSim frame	594,643	Transactions in the source file
Fraudulent transactions retained	7,200	All transactions with a fraud label of 1
Nonfraudulent transactions retained	10,000	Random sample from 587,443 transactions with a fraud label of 0
Nonfraudulent transactions not selected	577,443	Excluded by the sampling procedure
Constructed analytical sample	17,200	Combined retained transactions
Final estimation sample	Not yet determined	Subject to the requirements of the selected estimator

The sampling procedure increased the fraud share from approximately 1.21 percent in the source frame to 41.86 percent in the analytical sample. The workbook provides sampling weights to account for this selection when estimating quantities for the source frame.

3.2 Data Integration, Cleaning, and Quality Assurance:

The workbook was constructed from one transaction file. Each selected record retains its source file row number, simulation step, customer identifier, and merchant identifier. The 17,200 source row numbers are unique. Repeated customer and merchant identifiers represent transactions linked to the same simulated actors and were retained.

The selected variables contain no blank cells. The age code U appears in 29 transactions and is retained as an original category whose meaning requires verification against the source documentation. The sample contains one transaction with an amount of zero and no negative amounts. No currency conversion, inflation adjustment, interpolation, or imputation was applied. No transaction was removed as an outlier.

Transaction amounts range from zero to 8,329.96. The available file does not establish whether an extreme amount is an error, so the original values remain unchanged. The source describes amounts as monetary transaction values, but the currency designation should be confirmed before currency specific interpretation. Jha

and Westland 2013 show that examining transaction patterns can support fraud analysis, while the interpretation of individual values depends on their transaction context.

The category distribution creates an estimation constraint. Some payment categories contain only one fraud class in the constructed sample. These categories can cause complete separation in an ordinary logistic regression. Category definitions and treatment of separation must therefore be specified before estimation.

3.3 Variable Measurement, Construction, and Validation:

The dependent variable is the source fraud label. The three explanatory variables are transaction amount, payment category, and customer age code. Each is an observed field retained from BankSim. No latent construct or composite index was constructed.

Variable	Analytical Role	Definition and Construction	Unit and Coverage	Treatment and Limitation
Fraud label	Outcome	Source label identifying a transaction as fraudulent or nonfraudulent	One label per transaction; 17,200 valid records	Zero indicates nonfraudulent; one indicates fraudulent
Transaction amount	Predictor	Monetary value recorded for the transaction	Monetary units per transaction; 17,200 valid records	Original values retained; currency designation requires confirmation
Payment category	Predictor	Source category assigned to the payment	Categorical value per transaction; 15 categories	Original categories retained; some contain only one fraud class
Customer age code	Predictor	Source age category assigned to the customer	Categorical value per transaction; eight recorded codes	Original codes retained; the meaning of U requires verification
Sampling weight	Sample adjustment	Inverse of the selection probability within each fraud class	One weight per transaction	Fraud weight is 1; nonfraud weight is 58.7443

The fraud label is the simulated outcome, not an estimate of detection accuracy. Transaction amount is a continuous measure. Payment category and age code are categorical measures and require categorical coding for estimation. The age codes are not treated as a continuous age scale.

Descriptive Measure	Valid Transactions	Mean	Standard Deviation	Median	Minimum	Maximum
Transaction amount	17,200	240.86	594.41	46.72	0.00	8,329.96

These statistics describe the constructed sample without applying sampling weights. The source fraud label contains 7,200 fraudulent and 10,000 nonfraudulent transactions. The imbalance between source classes and the altered sample composition matter when interpreting fraud probabilities. Dal Pozzolo et al. 2018 identify class imbalance as a central issue in transaction fraud analysis.

3.4 Analytical Framework, Identification, and Reproducibility:

The intended analysis is associational. It examines whether transaction amount is associated with recorded fraud status and whether fraud status differs across payment categories and customer age codes. The estimated relationships would describe patterns in the simulated BankSim transactions. They would not identify causal effects.

A binary outcome model can address the three questions using the recorded fraud label and the three observed predictors. Payment category and age code require reference categories. The choice of estimator must address categories with complete separation, and uncertainty estimates should account for repeated transactions associated with customers where the estimator permits. No baseline model or inferential procedure has yet been implemented, so model coefficients, estimation losses, and hypothesis outcomes are not reported.

Sampling weights support estimates for the source transaction frame when applied consistently with the estimation method. If a predictive model is developed, evaluation should account for the original fraud prevalence. Saito and Rehmsmeier 2015 explain why precision and recall are relevant when evaluating classifiers with imbalanced outcomes. Randhawa et al. 2018 also demonstrate that fraud detection performance depends on the model and data used for evaluation.

The sample can be reproduced from the source file by retaining all fraudulent transactions and drawing 10,000 nonfraudulent transactions with the random seed recorded in the workbook. The original source row number allows each selected transaction to be traced to its source record. Reproducibility also depends on retaining the same source file version and sampling software.

4. Methods:

4.1 Research Design and Identification:

The analysis estimates conditional associations between transaction characteristics and recorded fraud status in simulated bank payments. Each observation represents one transaction. Variation in transaction amount,

payment category, and customer age code supplies the comparisons used to examine H1 and the two research questions.

The outcome is the binary fraud label assigned in BankSim. The analysis does not evaluate a treatment or identify an external source of variation. Its estimates therefore describe associations within the simulated source frame. They do not establish that a transaction characteristic causes fraud or that the relationships apply to actual banking transactions. Stojanović et al. 2021 describe BankSim as a simulated payment dataset that includes transaction characteristics and fraud labels.

The sample construction table and variable dictionary in the data section document the source, selection rules, and measures. The main estimation sample contains 17,200 transactions associated with 3,927 customer identifiers. All model variables are observed for these transactions.

4.2 Model Specification and Measurement:

The main specification is a weighted linear probability model:

$$Y_i = \alpha + \beta A_i + \sum_{j=1}^{14} \gamma_j C_{ij} + \sum_{k=1}^{7} \delta_k G_{ik} + \varepsilon_i$$

In Equation 1, Y_i is the fraud label for transaction i, coded one for fraud and zero for nonfraud. A_i is the transaction amount divided by 100. C_{ij} represents the payment category indicators, and G_{ik} represents the customer age code indicators. The term $\alpha$is the intercept, and $\varepsilon_i$is the model residual.

The amount coefficient βmeasures the fitted change in fraud probability associated with an increase of 100 monetary units, conditional on the included category and age indicators. Payment category and age code enter as categorical variables. The reference payment category is bars and restaurants, and the reference age code is zero. The model includes no interaction, lag, or fixed effect.

The fraud label and predictors are observed fields rather than latent constructs. The model treats the source age codes as categories rather than numerical ages. The code U is retained as a separate category because its meaning requires verification against the source documentation.

4.3 Estimation and Inference:

The analysis estimates Equation 1 by weighted least squares. The weights reflect the sample selection described in the data section. Every fraudulent transaction receives a weight of one. Each sampled nonfraudulent transaction receives a weight of 58.7443, calculated as the number of nonfraudulent transactions in the accessible source frame divided by the number selected.

The weights are treated as fixed selection weights in estimation. They adjust the relative contribution of the two fraud classes toward the source transaction frame. The resulting estimates remain conditional on the selected source file and its simulated labels.

The analysis calculates customer clustered standard errors because multiple transactions can share a customer identifier. The covariance calculation permits heterogeneity and dependence within customer clusters and uses 3,927 clusters. A finite sample correction is applied using the number of observations, estimated coefficients, and clusters. This procedure does not account for every possible dependence across customers, including dependence associated with shared merchants or simulation steps. Cameron and Miller 2015 explain the purpose and assumptions of clustered inference.

The amount coefficient is assessed using a normal reference distribution and a 95 percent confidence interval. Joint Wald tests, evaluated against chi square reference distributions, assess the payment category indicators and the age code indicators. The main model uses all 17,200 records. No missing observations were removed, and no imputation or outlier exclusion was applied.

Estimation was performed in Python 3.12.14 using NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0, and openpyxl 3.1.5. The weighted coefficients were calculated from the weighted least squares normal equations. The sampling procedure used the random seed recorded in the workbook.

4.4 Sensitivity, Model Limits, and Reproducibility:

Two exploratory sensitivity analyses assess the amount association under different sample conditions. The first estimates the same specification without sampling weights. The second retains only payment categories containing both fraudulent and nonfraudulent transactions and estimates the weighted model on 7,734 records. The latter changes the analytical sample as well as the category coverage, so its coefficient is interpreted as a comparison for the restricted sample.

The analysis checks whether fitted values fall within the zero to one range. The main model produces fitted values outside that range for 5,471 transactions. Its coefficients are consequently interpreted as linear association summaries, not as reliable individual fraud probabilities. Some payment categories also contain only one fraud class in the constructed sample, which limits comparisons involving those categories.

The workbook retains source row numbers and the sampling weights needed to trace selected records and reproduce the estimation sample. Reproduction of the estimates requires the same workbook, categorical reference levels, weighting convention, customer clustering, and software settings described here. The

simulation, selected sample, and linear specification limit transfer to other payment systems and do not provide causal identification or an assessment of operational detection performance.

5. Findings:

This analysis examines whether transaction amount, payment category, and customer age code are associated with recorded fraud status. It evaluates H1 and addresses the two research questions using simulated bank payment records. The estimates describe associations within the BankSim source dataset.

The analytical sample comprises 17,200 transactions, including 7,200 fraudulent and 10,000 nonfraudulent payments. These records are associated with 3,927 customer identifiers and cover 180 simulation steps. Every transaction has a recorded fraud label, amount, payment category, and customer age code. Because the sample retains all fraudulent transactions but only a subset of nonfraudulent transactions, the analysis applies the supplied sampling weights when estimating relationships for the source transaction frame. Dal Pozzolo et al. 2018 emphasize that class imbalance and the conditions under which transaction data are evaluated affect the interpretation of fraud detection results.

The main analysis uses a weighted linear probability model with fraud status as the outcome. Transaction amount enters as a continuous predictor, while payment category and age code enter as categorical predictors. Standard errors are clustered by customer to account for dependence among transactions associated with the same customer. The amount coefficient is expressed as the change in fitted fraud probability, in percentage points, for an increase of 100 monetary units. Joint tests assess whether the category coefficients and age code coefficients differ from zero as groups.

Relationship	Adjusted Estimate or Joint Test	Statistical Result
Transaction amount	4.27 percentage points per 100 monetary units	95 percent confidence interval from 3.88 to 4.65; $p < 0.001$
Payment category	Wald statistic of 59,409.20 with 14 degrees of freedom	$p < 0.001$
Customer age code	Wald statistic of 2.40 with 7 degrees of freedom	$p = 0.934$

Transaction amount has a positive adjusted association with recorded fraud status. Holding payment category and customer age code constant, an increase of 100 monetary units is associated with a 4.27 percentage point increase in fitted fraud probability. The confidence interval extends from 3.88 to 4.65 percentage points and excludes zero. The evidence supports H1 within the specified model. This is a conditional association for the simulated transactions; it does not establish that increasing a payment amount causes fraud.

The difference between fraud classes is also visible in the unadjusted amount distribution. Fraudulent payments have a median amount of 319.18 monetary units, while nonfraudulent payments have a median of 26.64. The corresponding means are 530.93 and 32.02 monetary units. The gap between the mean and median among fraudulent payments indicates that some large transactions raise the average above the median. Jha and Westland 2013 demonstrate the value of examining transaction patterns when studying card fraud, while Correa Bahnsen et al. 2016 show that the way transaction characteristics are represented can affect fraud detection analysis.

Payment category is jointly associated with recorded fraud status after adjustment for transaction amount and customer age code. The joint test rejects the hypothesis that all category coefficients are zero, with $p < 0.001$. The answer to RQ1 is therefore that recorded fraud status differs across payment categories in the fitted model. The joint result identifies variation across the category indicators; it does not by itself establish that every individual category differs from every other category.

The weighted fraud shares illustrate the extent of variation in the constructed sample. The estimated share is 9.50 percent for health payments, 38.34 percent for hotel services, and 42.31 percent for sports and toys. No sampled transportation payments have a fraud label, while all sampled leisure payments have a fraud label. These shares are unadjusted category summaries, whereas the joint test also accounts for transaction amount and age code. The contrast between the two types of evidence matters because payment categories may have different amount distributions.

The extreme category patterns also limit interpretation. Several categories contain only one fraud class in the constructed sample, and the nonfraudulent transactions were sampled from a larger frame. Weighted category shares are estimates for that frame and remain subject to sampling uncertainty. The patterns also reflect a simulator that combines ordinary payments with specified fraudulent behavior. Stojanović et al. 2021 examine BankSim alongside other fraud datasets and show that detection results vary across data and methods. The category differences observed here should therefore be interpreted as properties of this simulated dataset rather than universal category risks.

Customer age code is not jointly associated with fraud status after adjustment for amount and payment category. The joint Wald statistic is 2.40 with seven degrees of freedom, and $p = 0.934$. For RQ2, the analysis does not find evidence of an adjusted difference across the recorded age categories. This result does not prove

that all age groups have equal fraud rates. It indicates that the fitted model does not detect a difference using the available age codes and sample. The code U appears in 29 transactions, and its substantive meaning requires verification before it can be interpreted as an age group.

An exploratory sensitivity analysis retained only payment categories containing both fraud and nonfraud observations. This restriction reduced the sample to 7,734 transactions. The transaction amount coefficient remained positive at 4.51 percentage points per 100 monetary units, with a 95 percent confidence interval from 4.08 to 4.94 percentage points. Its direction is consistent with the full sample estimate, although the restricted analysis represents a different set of payment categories. The comparison supports the stability of the amount association within categories where both fraud classes are observed, but it does not resolve the uncertainty associated with the excluded categories.

A second exploratory comparison estimated the model without sampling weights. The amount coefficient remained positive but was smaller, at 0.78 percentage points per 100 monetary units. This difference should not be read as a direct test of which estimate is correct. The weighted and unweighted analyses assign different influence to the two fraud classes and therefore describe different transaction compositions. Bhattacharyya et al. 2011 demonstrate that fraud detection results depend on analytical choices and data characteristics, reinforcing the need to report the sampling design with the estimates.

The model has a material limitation for prediction. It generated fitted values outside the zero to one range for 5,471 transactions. The coefficients can be read as summaries of linear associations, but those fitted values cannot be interpreted as valid individual probabilities. The analysis has not evaluated a fraud classification rule, detection threshold, precision, or recall. Saito and Rehmsmeier 2015 explain why precision and recall are useful when assessing classification performance under class imbalance. Those measures would require a separate predictive evaluation.

The findings support H1 within the weighted linear model and show differences in recorded fraud status across payment categories. They do not provide evidence of an adjusted association for customer age code. The amount result remains positive in the restricted category analysis, but the simulated setting, category separation, sampling design, and out of range fitted values limit broader interpretation. The findings describe the supplied BankSim transactions and do not establish causal effects or detection performance in an operating bank.

6. Discussion:

Financial fraud analysis gains meaning when transaction characteristics are interpreted together rather than treated as independent warning signals. The weighted model indicates that larger transaction amounts are associated with a higher likelihood of a simulated fraud label after accounting for payment category and customer age code. Payment categories also differ jointly, while the age codes do not show an adjusted association. This combination sharpens the interpretation of H1: the amount relationship persists within the specified model, but it does not establish that amount causes fraud. The contribution is a joint, sampling weighted assessment of these three characteristics within the BankSim transactions. It does not establish that this combination has never been studied. Al-Hashedi and Magalingam2021 show that financial fraud research spans different data and detection methods, making the analytical setting essential to interpreting any reported relationship.

The distinction between transaction amount and payment category suggests that transaction value should be interpreted in context. The positive amount coefficient remains when category is included and also remains positive in the exploratory sample restricted to categories containing both fraud classes. This pattern is consistent with an amount association that is not explained entirely by the observed category indicators. It does not identify how or why the association arises, and the analysis does not establish a mediating pathway. Van Vlasselaer et al. 2015 combine transaction characteristics with customer and merchant network information, illustrating that payment context can extend beyond the individual fields examined here. The present results identify a reason to investigate such context, but they do not measure the additional information supplied by a network model.

The absence of a detectable joint age code association adds a boundary to the transaction anomaly framework. The result suggests that, after accounting for amount and payment category in this model, the recorded age categories do not provide evidence of a further association with the fraud label. It does not demonstrate that age is irrelevant to fraud in other datasets or that customers across age groups face equal risks. The codes describe simulated customers, and the meaning of one code remains unverified. This distinction matters because transaction characteristics and customer characteristics can represent different kinds of information. Jha, Guillen, and Westland 2012 examine aggregated transaction behavior, providing a basis for asking whether information about spending patterns adds value beyond the individual transaction fields used in this analysis.

The category results expose a constraint in interpreting simulated fraud patterns. Some sampled categories contain only fraudulent or only nonfraudulent records, producing strong contrasts that may reflect both the simulation design and the selection of nonfraud records. The weighted and unweighted amount

estimates also differ, showing that the choice of transaction composition affects the reported association. These patterns make sampling and model specification part of the substantive interpretation rather than merely technical details. The linear probability model provides an adjusted association, but its out of range fitted values limit its use for individual probability estimation. Fiore et al. 2019 address class imbalance through a different classification approach, underscoring that the handling of fraud and nonfraud observations must be considered alongside the intended analytical outcome.

The international relevance lies in the distinction between identifying associations in a public simulation and demonstrating performance in operational payment systems. The BankSim evidence suggests that amount and payment category merit examination together, while the age result cautions against assuming that every available customer characteristic adds explanatory value. For practitioners, the immediate implication is to evaluate these fields against locally observed transactions before using them to set alerts or make decisions. A further analysis could test whether the amount association remains stable in real payment data and whether category patterns persist when both fraud classes are well represented. Dal Pozzolo et al. 2014 describe fraud detection as an operational task shaped by changing transaction streams and investigator requirements. That setting provides a standard for future external validation, rather than evidence that the present model is ready for deployment.

7. Conclusion and Implications:

Fraud detection systems face a common challenge: a transaction can appear unusual without being fraudulent. This analysis shows that payment characteristics provide different information when considered together. The positive association between payment value and recorded fraud remains after adjustment for transaction context and customer characteristics, while the evidence does not establish an additional association for the recorded age categories. The contribution is a weighted assessment of these relationships in a simulated payment system where fraudulent and legitimate records were sampled at different rates. It refines transaction anomaly reasoning by showing that the interpretation of a signal depends on the accompanying transaction context and the composition of the data. These relationships are associations and do not identify causal pathways. For financial institutions, the findings support assessing transaction signals jointly rather than applying a single characteristic as a fraud rule. Risk teams should also account for sampling when interpreting estimated fraud patterns. For regulators and system designers, the results support requiring evidence that proposed detection tools perform adequately on representative transaction data before operational use. Better evaluation can help institutions direct investigations while limiting unnecessary disruption to legitimate customers.

The findings provide a basis for further validation across payment environments. The data are simulated, some payment categories contain only one fraud class, and the fitted model produces values outside the probability range for some transactions. Future research can examine whether the observed associations persist in representative banking data, use models suited to bounded probabilities, and assess detection performance under realistic fraud prevalence. Comparisons across institutions and payment systems could establish whether category patterns transfer across settings. Additional research could also examine transaction histories and customer or merchant relationships to determine whether they improve identification beyond the characteristics assessed here.

References:

1. Al-Hashedi, K. G., & Magalingam, P. 2021. Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 40, Article 100402. https://doi.org/10.1016/j.cosrev.2021.100402
2. Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. 2011. Data mining for credit card fraud: A comparative study. Decision Support Systems, 50, 602-613. https://doi.org/10.1016/j.dss.2010.08.008
3. Cameron, A. C., & Miller, D. L. 2015. A practitioner's guide to cluster robust inference. Journal of Human Resources, 50, 317-372. https://doi.org/10.3368/jhr.50.2.317
4. Correa Bahnsen, A., Aouada, D., Stojanovic, A., &Ottersten, B. 2016. Feature engineering strategies for credit card fraud detection. Expert Systems with Applications, 51, 134-142. https://doi.org/10.1016/j.eswa.2015.12.030
5. Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. 2018. Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29, 3784-3797. https://doi.org/10.1109/TNNLS.2017.2736643
6. Dal Pozzolo, A., Caelen, O., Le Borgne, Y. A., Waterschoot, S., & Bontempi, G. 2014. Learned lessons in credit card fraud detection from a practitioner perspective. Expert Systems with Applications, 41, 4915-4928. https://doi.org/10.1016/j.eswa.2014.02.026
7. Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. 2019. Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences, 479, 448-455. https://doi.org/10.1016/j.ins.2017.12.030

8. Jha, S., Guillen, M., & Westland, J. C. 2012. Employing transaction aggregation strategy to detect credit card fraud. Expert Systems with Applications, 39, 12650-12657. https://doi.org/10.1016/j.eswa.2012.05.018
9. Jha, S., & Westland, J. C. 2013. A descriptive study of credit card fraud pattern. Global Business Review, 14, 373-384. https://doi.org/10.1177/0972150913494713
10. Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. 2018. Credit card fraud detection using AdaBoost and majority voting. IEEE Access, 6, 14277-14284. https://doi.org/10.1109/ACCESS.2018.2806420
11. Saito, T., &Rehmsmeier, M. 2015. The precision recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE, 10, Article e0118432. https://doi.org/10.1371/journal.pone.0118432
12. Seera, M., Lim, C. P., Kumar, A., Dhamotharan, L., & Tan, K. H. 2021. An intelligent payment card fraud detection system. Annals of Operations Research. Advance online publication. https://doi.org/10.1007/s10479-021-04149-2
13. Stojanović, B., Božić, J., Hofer-Schmitz, K., Nahrgang, K., Weber, A., Badii, A., Sundaram, M., Jordan, E., & Runevic, J. 2021. Follow the trail: Machine learning for fraud detection in fintech applications. Sensors, 21, Article 1594. https://doi.org/10.3390/s21051594
14. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. 2015. APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. Decision Support Systems, 75, 38-48. https://doi.org/10.1016/j.dss.2015.04.013