

Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2024 Proceedings

International Conference on Information
Systems (ICIS)

December 2024

Artificial Intelligence to Improve Public Budgeting

Dominic Santschi

University of St. Gallen, dominic.santschi@unisg.ch

Marc Christopher Grau

University of St. Gallen, marcchristopher.grau@unisg.ch

Dennis Fehrenbacher

University of St. Gallen, dennis.fehrenbacher@unisg.ch

Ivo Blohm

University of St. Gallen, ivo.blohm@unisg.ch

Follow this and additional works at: <https://aisel.aisnet.org/icis2024>

Recommended Citation

Santschi, Dominic; Grau, Marc Christopher; Fehrenbacher, Dennis; and Blohm, Ivo, "Artificial Intelligence to Improve Public Budgeting" (2024). *ICIS 2024 Proceedings*. 1.

https://aisel.aisnet.org/icis2024/iot_smartcity/iot_smartcity/1

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2024 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Artificial Intelligence to Improve Public Budgeting

Short Paper

Dominic Santschi

University of St.Gallen
St.Gallen, Switzerland
dominic.santschi@unisg.ch

Marc C. Grau

University of St.Gallen
St.Gallen, Switzerland
marcchristopher.grau@unisg.ch

Dennis D. Fehrenbacher

University of St.Gallen
St.Gallen, Switzerland
dennis.fehrenbacher@unisg.ch

Ivo Blohm

University of St.Gallen
St.Gallen, Switzerland
ivo.blohm@unisg.ch

Abstract

Accurate public budgeting is essential for efficient resource allocation and societal trust. However, recent studies have shown that public budgets often project deficits but have substantial surpluses. This budgetary slack can lead taxpayers to overpay for services not rendered, delay necessary investments, or distort public perceptions of government efficiency. To avoid such unfortunate outcomes, we study how artificial intelligence (AI) can help decision-makers in the public sector with budgeting. We operationalize our research question using a two-step approach. First, we utilize open data from Swiss financial authorities to train and test an AI model. Our preliminary results validate the potential of AI to predict public accounts better than human experts. Second, we study whether decision-makers effectively utilize the AI model in an experimental scenario. The results of our experimental study indicate that human-AI collaboration could indeed support decision-makers to improve public budgeting by reducing budgetary slack.

Keywords: Open Data Utilization, Human-AI Collaboration, Data-driven Governance, Budgeting

Introduction

Budgeting is an integral part of the management control system in organizations (Datar & Rajan, 2021). However, compared to private sector organizations, public sector organizations have received relatively little research on their budgeting practices and potential improvements (Aranda et al., 2023). This is despite the public sector's unique complexities of managerial reforms and promising emerging opportunities for processual improvements (Anessi-Pessina et al., 2016; Sicilia & Steccolini, 2017). For the public sector, accurate budgeting is essential for fiscal responsibility and efficient resource allocation (Valle-Cruz et al., 2022). Overestimated deficits or underestimated surpluses can signal false financial distress, influencing controversial policy decisions such as austerity or resistance to tax reductions (Marlowe, 2011). Administrative divisions of the public sector have been found to project significant deficits but have substantial surpluses, undermining trust in the budgeting process and suggesting a lack of fiscal discipline (Rühli & Hutter, 2023). This budgetary slack (Anthony & Govindarajan, 2007) can lead taxpayers to overpay for services not rendered or delay necessary investments, distorting public perceptions of government efficiency and eroding trust in the public sector.

To avoid such unfortunate outcomes, organizations explore collaborations between humans and technology (Choudhary et al., 2023; Valle-Cruz et al., 2022). AI systems can create predictions that then can undergo human judgment such that human decision-makers can follow, adapt, or overrule AI-based forecasts (Kaplan & Haenlein, 2019). This process of human-AI collaboration affords fine-grained socio-technical arrangements to combine the strengths of humans, e.g., intuition, and AI systems (Antretter et al., 2020).

Such a configuration can further improve the quality and reliability of accounting estimates (Bertomeu, 2020) and could avoid dysfunctional consequences provoked by managerial manipulations (Jensen, 2003). Yet, human decision-makers often hesitate to use AI predictions. They do not understand how such predictions are created (Huysmans et al., 2011), have little trust in their accuracy (Castelo et al. 2019), and penalize algorithm prediction errors stronger than human ones (Dietvorst et al., 2015). This well-documented behavioral phenomenon of algorithm aversion (Dietvorst et al., 2015; 2018) hinders the effective implementation of AI in budgeting. We seek to contribute to theory and support budgeting practitioners in the public sector by answering the following research question:

How can AI help decision-makers in the public sector with budgeting?

We respond to our research question with a two-step approach. First, we study how well AI can predict public accounts in comparison with human experts. For this purpose, we developed and tested an AI model based on actual expense data derived from an open dataset gathered from the financial authorities of Switzerland (FDK-SDK, 2024). Our preliminary results validate the potential of an AI model to predict public accounts better than humans. Building on this foundation, it is crucial to determine whether the model's demonstrated superior performance is being effectively utilized by decision-makers in practice. The second step of our research addresses this by exploring how an AI slack evaluation translates into practical advantages for decision-makers, specifically in reducing budgetary slack. To answer this question, we conducted an initial experiment with human experts with proven budgeting experience. Actual investment expense data informed the experimental instrument on public sector budgeting. The initial analysis of the budgets and the sentiment analyses of the textual justifications reveal that AI could indeed improve public budgeting by decreasing budgetary slack.

Background

Related Work

Academic research has extensively explored the concept of collaboration in budgeting, with a particular focus on the role of negotiation. For instance, Fisher et al. (2000) find that negotiations between lower-level and upper-level management on budgets can help to reduce budgetary slack. Yet, if negotiations fail and upper-level management enforces a budget, motivation is reduced. Arnold (2015) studies the effect of exogenous constraints on budget negotiations. He provides evidence that financial pressure on upper-level management and increased opportunity costs lead to more cooperative negotiation behavior. He adds to Fisher et al.'s (2000) finding that under such constraints, upper-level management's imposition of a budget does have a weaker negative motivational effect. Brownell (1982) demonstrates that lower-level management participation in the budgeting process increases managerial performance conditional on the individual characteristic locus of control. Arnold and Gillenkirch (2015) provide evidence that lower-level management cooperation and performance increase when the same budget figures are used for planning, target-setting, and related performance evaluation purposes. A common perspective among these studies is the notion that budgeting is a process that involves people.

However, this research has largely not considered the possible involvement of AI and the potential benefits of human-AI-collaboration in the public sector (Di Vaio et al., 2022). While accounting academics are starting to show a strong interest in the way decision-makers use AI systems, the literature is in its early stages (Commerford et al., 2022; Chen et al. 2022, Fehrenbacher et al. 2023). The majority of these cases focus on the behavioral phenomenon of algorithm aversion (Dietvorst et al., 2015; 2018), which hinders the simple and effective implementation of AI in budgeting. Despite AI's potential to improve budgeting, human decision-makers are frequently reluctant to use this emerging technology. They do not understand how predictions are created (Huysmans et al., 2011), have little trust in their accuracy (Castelo et al. 2019), and penalize algorithm prediction errors stronger than human ones (Dietvorst et al., 2015). Chen et al. (2022) provide evidence that managers perceive data analytics as less credible than human experts when advice entails bad news. Similarly, Commerford et al. (2022) find that auditors propose smaller adjustments when advice comes from an AI tool rather than a human expert. Fehrenbacher et al. (2023) document that discounting algorithmic advice is conditional to the data the algorithm used, whereas more complex and less structured data increases advice discounting. In contrast to these hypothetical cases explored in the current accounting literature, we use a purpose-built AI model and bring the business and accounting community one step closer to understanding the potential of AI in budgeting.

Hypothesis Development

AI vs. Human Experts: Who Budgets Better?

Budgetary slack is a widely recognized problem in budgeting and target-setting processes that can reduce individual and organizational performance (Chow et al., 1991), yet it is not considered as being an illegal or fraudulent activity to the extent that the given organizational leeway is used to move organizational budgets in a desired direction (Datar & Rajan, 2021). We define budgetary slack as the practice of setting a revenue budget lower than the “best estimate” available at a given time and an expense budget higher than the “best estimate” (Anthony & Govindarajan, 2007).

Agency theory offers further insights into the underlying reasons for the differences between AI and human experts. According to agency theory (Jensen & Meckling, 1976), principals (e.g. higher-level managers) frequently rely on agents (e.g. lower-level managers) “best estimates” to form opinions about the financial performance in the upcoming period. Such participation needs open communication about the state of a business from agents to principals. However, the “best estimates” of agents at a given time are hard to tell for principals who oversee the budgeting process. Agency theory assumes that agents are self-interested, work-averse, and risk-averse. These assumptions can lead to misalignments in budgeting (Covaleski et al., 2003). Self-interest and work-aversion of agents can lead to budgetary slack because targets may be achieved with less effort and a higher likelihood of possibly gaining financial reward when budgetary slack is built in. For instance, although an agent estimates investment expenses for the upcoming year at 1,000 monetary units (“best estimate”), they may budget a safety buffer of 10% more to increase the likelihood of performing better than budgeted. Risk aversion can lead to budgetary slack as agents may view slack as a safety buffer for unexpected unfavorable events (Dunk, 1995). While agency theory may significantly influence human experts’ behavior, these dynamics generally do not directly affect AI models (Blohm et al., 2022).

Perceptual Dynamics of AI

Agency theory suggests that an effective solution to budgetary slack lies in narrowing the information gap between agents and principals, thereby better aligning the interests of agents with those of the principals and enhancing the budgeting process (Liessem et al., 2015). A common approach is to examine the roles in the budgeting process, typically those of lower-level and upper-level management, to improve budgetary outcomes (Van der Stede, 2000). The notion in this research is that cooperation of various levels in budgeting (e.g., through elements of participatory budgeting and negotiations) can lead to improvement in motivation and performance (Fisher et al., 2000).

The advice-taking literature has established task-related competence as a key factor that influences taking advice (Bonaccio & Dalal, 2006). The human-computer interaction literature has documented some evidence that AI competence influences users’ choice of an AI system (Gilad et al., 2021). In the context of budgeting the portrayal of task-related competence involves an AI’s ‘understanding’ of budgetary slack. We expect that an AI can demonstrate higher levels of task-related budgeting competence in not only making its budgeting prediction but also making a specific evaluation on the creation of budgetary slack. An AI portraying task-related competence on budgetary slack should then influence decision-makers in reducing slack. To do so, the AI would need to distinguish between a “best estimate” and the possible inclusion of budgetary slack to evaluate whether a decision maker’s prediction possibly involves budgetary slack. We expect that if such an evaluation takes place it increases individuals’ perception that the AI ‘understands’ the context of the budgeting task and is competent. We call such an evaluation “AI slack evaluation”. This stands in contrast to an AI ‘just’ predicting a possible future budget figure. In analogy with agency theory (Jensen & Meckling, 1976; Liessem et al., 2015), we expect that an appropriate AI slack prediction can narrow the information gap in a budgeting context and, as such, can limit budgetary slack. If an AI predicts slack, it is perceived to possess higher budgeting competence. As we will outline, we aim to base our evidence on the outcomes of purpose-built AI models that improve public budgeting. Thus, we state our main hypothesis:

Main Hypothesis (H) 1. *An AI slack evaluation (vs. no AI slack evaluation) decreases budgetary slack in public budgeting.*

Development and Evaluation of AI Model for Public Budgets

Similar to the private sector, the public sector frequently provides access to fine-grained budgeted and actual financial data. For instance, in Switzerland, the Federal Statistical Office houses The Open Government Data Office (see: opendata.swiss). However, such open data initiatives have not been used extensively in research and practice besides their promising potential (Quarati, 2023). Thus, the non-confidential nature of this data allows us to build an AI model for public budgeting. Further, this data access helps us to study how public budgeting practices can be improved, which is of high societal relevance.

For the development and evaluation of our AI model, we used publicly available open data for a total of 36 income and expense accounts from 26 Swiss cantons – the main administrative divisions of the Swiss public sector (FDK-SDK, 2024). We trained models relying on different versions of gradient tree boosting algorithms (He et al., 2019) including XGBoost (Chen & Guestrin, 2016) and CatBoost (Dorogush et al., 2018). For the training procedure, we used historical data from the years 2011 to 2021. To validate our model predictions we used the year 2022 as unseen testing data. Thus, the model never sees more data than any human expert would to make the predictions for the testing period. We evaluated the AI model performance by comparing the difference between the model budget forecast and the actual performance of 2022 (i.e., the canton's final income and spending in 2022). To test who (our AI model vs. human budgeting experts) predicts the actual performance more accurately, we compared the canton's official budget forecast from 2022 against our AI-based budget forecast. For this comparison, we calculated the Root Mean Squared Error (RMSE).

Overall, the initial results of our AI model are promising. Table 1 shows the aggregated results for all the income and expense accounts (n=36) reported by the cantons. In three out of four account groups, our AI model outperformed its human counterparts. Yet, we have to acknowledge that the forecasting performance varies across budgeting categories and that we have to engage in further research to make the forecasts more reliable and applicable to more granular budgeting categories.

Account Groups	# Accounts	AI Model Forecast	Human Forecast	Relative Difference
Expenses	9	106,400	91,517	+16%
Income	10	123,247	142,061	-13%
Investment Expense	8	23,515	36,259	-35%
Investment Income	9	19,669	20,259	-3%

Table 1. RMSE of our AI model and human experts for 2022 (in ,000 Swiss Francs)

Experimental Test of AI Model for Public Budgets

Experimental Design and Participants

To test our main hypothesis, we built upon our AI model to inform an experimental instrument for the public budgeting context. We conducted a 2x1 between-participant experiment with 125 participants who possess budgeting knowledge. Participants were randomly assigned to one of the two conditions of our treatment variable AI slack evaluation: (1) message with AI slack evaluation and (2) message without AI slack evaluation. Budgeting knowledge was assessed in a pre-study using several questions. We considered this important as contextual knowledge of budgeting was essential to complete the budgeting task. If pre-study participants failed to respond to one of the questions correctly, participants were not admitted to our study and we did not collect study data from those participants. Participants were recruited from the online panel Prolific and received a base compensation for participation in our study and extra compensation to justify their budgeting decision. We rewarded participants with £1 for participation and a bonus payment of £0.30 for successful completion (median time = 7 min 42 sec). These payments were set to achieve a “good compensation level” that exceeds the minimum Prolific requirements. A confirmation of ethical soundness was obtained from the university where the experiment took place.

Experimental Instrument

The experimental task involved participants acting as the mayor of a hypothetical city. Participants were asked to make budgeting decisions on the city's investment expense account, which was derived from the open data, and provide a written justification for their budget. Figure 1 shows a summary of the scenario timeline that was shown to participants.

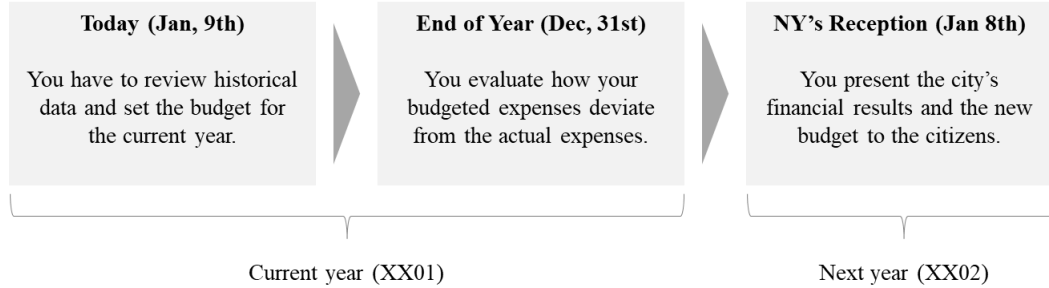


Figure 1. Scenario timeline provided to the participants

After participants submitted their initial budget, they received an evaluation message and were asked if they wanted to keep or revise their budget. This message contained the manipulation. Participants received either the *present* or *absent* condition of the AI slack evaluation treatment variable.

In the condition where the AI slack evaluation was *present*, participants were informed that the city's AI evaluated whether the budgeted expenses contained slack. When the budgeted expenses exceeded the prediction of our AI model, participants were informed that they planned slack according to the AI evaluation. Participants were also informed about the volume of slack according to the AI evaluation. In the *absent* condition, participants were reminded of their initial submission. In both conditions (*present* and *absent*) participants were reminded about possible negative effects if the budgeted expenses were too high and were asked whether they wanted to adjust their initial budgeting decision.

Sample and Manipulation Check

In our sample 62% identified as female, 37% as male, and 1% as other gender. 67% possessed more than 10 years of work experience and 71% were between 25 and 50 years of age. Our randomization procedure led to a balanced sample for gender, age, and work experience as we did not find statistically different distributions of gender, work experience, and age between our two conditions (all $p > .51$). To ensure the success of the manipulation, we asked participants in the post-experimental questionnaire to recall whether they had received an AI evaluation of their budget or not. All the participants of our final sample ($n=125$) answered the manipulation check correctly.

Results from Experimental Study

To test our main hypothesis, we calculated a budgetary slack reduction variable that captures the difference between the budgetary slack of the initial budget submission and the final budget submission after the participants were shown the treatment. We calculated the budgetary slack of the initial budget submission as the difference between the budgeted expenses of participants and the AI prediction. The budgetary slack of the final budget submission was calculated as the difference between the budgeted expenses in the final budget submission after the participants were shown the treatment and the AI prediction. Consistent with the definition of budgetary slack by Anthony and Govindarajan (2007) we treat the AI prediction as the "best estimate". As such if participants budgeted expenses higher than the prediction of our AI model, we expect participants to have budgeted slack.

Our results suggest that when participants over-budget investment expenses, AI slack evaluation (vs. no AI slack evaluation) reduces budgetary slack. From the 120 participants that budgeted slack in the initial submission, 118 reduced the slack or kept the slack constant from the initial budget submission to the final budget submission (55 in the *present* condition and 63 in the *absent* condition). Since all data used in the case is real data from a canton's expense account (FDK-SDK, 2024), we can comment on the difference in

size in CHF million. Results show that the participants in the AI slack evaluation *present* condition reduced their budgetary slack on average by approximately CHF 3.71 million while those in the *absent* condition reduced their budgetary slack on average by approximately CHF 42.6 thousand. Accordingly, the participants in the AI slack evaluation *absent* condition reduced their budgetary slack on average by approximately 0.02 percent while those in the *present* condition reduced their budgetary slack on average by approximately 1.98 percent. A Welch's t-test shows that the difference in budgetary slack reduction between the *present* condition (mean = -3,709.51, s = 8,823.05, n = 55) and *absent* condition (mean = -42.56, s = 337.77, n = 63) is significant ($t = -3.08$, $p = .002$).

Sentiment Analysis on Justifications

In addition to the budgets, we analyzed the written justifications collected from the participants of our experiment to justify their forecasting decisions. We thereby explored the potential of using textual analysis to understand the creation of budgetary slack in more detail. For instance, we studied how the presence of uncertainty, which is a possible determinant of budgetary slack (Dunk, 1995), can be captured by text mining techniques (Loughran & McDonald, 2011). Interestingly, our initial textual analysis of budgetary slack in the submissions of 120 participants reveals a predominance of uncertainty, polarity, and subjectivity. We explored the justifications provided by participants both before and following the evaluation message, where the AI was either *present* or *absent*. For instance, Responder-ID #130 stated in their initial budget submission: "I took an average of the Budgeting expenses over the 9 years and added an inflation factor of 5%. This yielded 203,177. Then, I did the same for the buffer zone (along with an inflation factor). This gave me 37,850. I added them together for the final budget. But this was greater than what could be selected on the sliding bar above. So I just picked the maximum." After the AI slack evaluation, their justification of the adjusted budget was "I trust the AI more than my calculations, so I gave myself a small buffer above what AI had come up with."

To inform our analysis, we employed two distinct methodologies. Firstly, we utilized the dictionary developed by Loughran and McDonald (2011) to assess the sentiment, particularly focusing on the uncertainty of the justifications. Secondly, we enriched our sentiment analysis with measures of polarity and subjectivity via natural language processing. Our preliminary observations suggest that participants inclined towards including budgetary slack in their submissions generally do so with neither clear positive nor negative sentiments (32 positive counts vs. 40 negative counts). Interestingly, our textual analysis also highlights a slight decrease in the uncertainty associated with justifications following the AI evaluation message, suggesting that the AI model's slack evaluation changed the tonality of the budget justifications. This finding could have important implications for budgeting practitioners, as the AI evaluation can eventually decrease uncertainty, which could ultimately be good for economic performance (Dunk, 1995).

Intended Contributions and Further Research

This study intends to make three contributions. We will discuss these contributions along with how we intend to shape this ongoing research - concerning AI model development, the design of the experimental scenario, and the application of textual analysis to budget justifications - to be able to provide more comprehensive evidence on our research question.

First, our AI model delivers initial evidence that it can budget public accounts more accurately than human experts. This contributes to the accounting field, which has traditionally neglected the interplay between budgetary slack and technological innovations in the public sector (Aranda et al., 2023). Also, our AI model contributes to the information systems literature by demonstrating a transferable application of technology-based organizational decision-making structures (Huysmans et al., 2011; Shrestha et al., 2019). To further strengthen our model, we need to augment our data to better account for exogenous factors. As of now, the data only consists of historical budgets. However, we see significant potential to enrich the current data with additional variables relevant to budgeting. These variables could include macroeconomic indicators such as the unemployment rate, population growth rate, growth rate of the gross domestic product, or inflation rate. As an alternative to macroeconomic variables, we could also review signals on tourism activity, political trends, and major events to name a few.

Second, our experimental study provides initial evidence that AI can help decision-makers improve budgeting. We plan to strengthen this contribution in various ways. First, we will test the prevalence of

budgetary slack in different scenarios. In the current experimental study, we only used one expense account of one Swiss canton and focused on participants who over-budget investment expenses (i.e., included a budgetary slack). Our limited focus on over-budgeting begs the question whether AI feedback is also effective in cases of under-budgeting. We will try to address this with future research to understand better what it is about AI advice in budgeting that makes decision makers re-decide their budgets. Furthermore, we plan to understand slack creation as influenced by AI for various accounts thereby aiming to classify revenue vs. expense accounts, contributing to the discussion initiated by Fahse et al. (2022), and discretionary vs. non-discretionary accounts. For this purpose, we are planning to use the corpus derived from the justifications in the experimental study to test our treatment against alternative control conditions. Thereby, we could strengthen the control condition and more directly test the different influences of an AI model versus a human expert evaluation. For the textual analysis, we aim to apply bidirectional encoder representations from transformers (BERT) (Devlin et al., 2019) to expand our current lexicon-based approach.

Third, our study makes an important contribution to data-driven governance and its budgeting processes in Switzerland and beyond. By developing systems that tackle budgetary slack, public budgeting experts can allocate resources more efficiently, leading to significant societal benefits. Public spending constitutes a substantial part of the gross domestic product in Western countries, underscoring the critical need to refine public budgeting practices. Improved budgetary processes could not only lead to more effective resource use but also indirectly reduce overall expenses, benefiting society at large. We plan to strengthen this contribution by challenging the ecological validity of our results with vetted budgeting experts in the public sector. To operationalize this complementary research, we plan to conduct vignette-based interviews with selected personnel thereby using the experimental scenarios on budgetary slack creation as vignettes. The vignette-based interview technique (Aguinis & Bradley, 2014) capitalizes on the public budgeting scenarios to help with a realistic evaluation of our AI interventions.

Conclusion

This research employs a two-step approach to study how AI can help decision-makers in the public sector with budgeting. In contrast to hypothetical cases explored in previous budgeting research, we use a specifically-trained AI model to increase our understanding of the potential of AI in budgeting. Our results suggest that when participants over-budget investment expenses, AI slack evaluation (vs. no AI slack evaluation) reduces budgetary slack. When there is no AI slack evaluation, we find that human experts are more prone to include slack in their budget, supporting the theoretical argument of inducing slack under uncertainty (Dunk, 1995) and aligning with our expectations from our historical data analysis of public sector budgeting data (FDK-SDK, 2024). By demonstrating task-related competence, we propose that an AI slack evaluation can improve budgetary estimates.

Participation and human-AI collaboration are essential to ensure compliance with societal and ethical expectations – especially because the successful adoption of AI projects in a public sector context depends on a complex set of factors and actors engaged (Puron-Cid, 2014). As AI continues to reshape the field, there is an urgent need for interdisciplinary discussion on potential risks and limitations (Eisikovits et al., 2024). Deploying AI agents to execute routine tasks such as budgeting in the management accounting discipline (Secinaro et al., 2024) presents a considerable risk to the traditional role of human employees, highlighting a critical area of concern (Busulwa & Evans, 2021).

This study develops and practically tests an AI model to enhance public budgeting by integrating information systems and behavioral research into accounting. Our AI model allows us to contribute to the management accounting literature (Aranda et al., 2023; Fisher et al., 2000), which has traditionally neglected the interplay between budgetary slack and technological innovations in the public sector. Also, our experimental study contributes to the behavioral sciences literature (Di Vaio et al., 2022; Puron-Cid, 2014; Shrestha et al., 2019; Valle-Cruz et al., 2022) by demonstrating a transferable application of technology-based organizational decision-making structures. Following our two-step research approach, we demonstrate AI's potential to enhance public sector budgeting by mitigating budgetary slack. We offer a novel approach to optimize resource allocation by integrating AI-generated estimates into organizational decision-making (Puranam, 2021). Given the substantial portion of GDP dedicated to public spending in Western countries, improving budgetary efficiency can indirectly reduce overall expenses, benefiting society.

References

- Aguinis, H., & Bradley, K. J. (2014). Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods*, 17(4), 351–371.
- Anessi-Pessina, E., Barbera, C., Sicilia, M., & Steccolini, I. (2016). Public sector budgeting: a European review of accounting and public management journals. *Accounting, Auditing & Accountability Journal*, 29(3), 491–519.
- Anthony, R. N., & Govindarajan, V. (2007). *Management Control Systems*, McGraw-Hill: Irwin.
- Antretter, T., Blohm, I., Siren, C., Grichnik, D., Malmström, M., & Wincent, J. (2020). Do Algorithms Make Better – and Fairer – Investments Than Angel Investors? *Harvard business review*.
- Aranda, C., Arellano, J., & Dávila, A. (2023). Budgeting in Public Organizations: The Influence of Managerial and Political Aspects. *European Accounting Review*, 32(2), 345–377.
- Arnold, M. C. (2015). The Effect of Superiors' Exogenous Constraints on Budget Negotiations. *The Accounting Review*, 90(1), 31–57.
- Arnold, M. C., & Gillenkirch, R. M. (2015). Using negotiated budgets for planning and performance evaluation: An experimental study. *Accounting, Organizations and Society*, 43, 1–16.
- Bertomeu, J. (2020). Machine learning improves accounting: discussion, implementation and research opportunities. *Review of Accounting Studies*, 25(3), 1135–1155.
- Blohm, I., Antretter, T., Sirén, C., Grichnik, D., & Wincent, J. (2022). It's a Peoples Game, Isn't It?! A Comparison Between the Investment Returns of Business Angels and Machine Learning Algorithms. *Entrepreneurship Theory and Practice*, 46(4), 1054–1091.
- Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, 101(2), 127–151.
- Brownell, P. (1982). The Role of Accounting Data in Performance Evaluation, Budgetary Participation, and Organizational Effectiveness. *Journal of Accounting Research*, 20(1), 12–27.
- Busulwa, R., & Evans, N. (2021). *Digital Transformation in Accounting (1st ed.)*. Routledge
- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, 56(5), 809–825.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785–794).
- Chen, C. X., Hudgins, R., & Wright, W. F. (2022). The Effect of Advice Valence on the Perceived Credibility of Data Analytics. *Journal of Management Accounting Research*, 34(2), 97–116
- Choudhary, V., Marchetti, A., Shrestha, Y. R., & Puranam, P. (2023). Human-AI Ensembles: When Can They Work? *Journal of Management*, 01492063231194968.
- Chow, C. W., Cooper, J. C., & Haddad, K. (1991). The Effects of Pay Schemes and Ratchets on Budgetary Slack and Performance: A Multiperiod Experiment. *Accounting, Organizations and Society*, 16(1), 47–60.
- Commerford, B. P., Dennis, S. A., Joe, J. R., & Ulla, J. W. (2022). Man versus Machine: Complex Estimates and Auditor Reliance on Artificial Intelligence. *Journal of Accounting Research*, 60(1), 171–201.
- Covaleski, M. A., Evans, J. H., Luft, J. L., & Shields, M. D. (2003). Budgeting Research: Three Theoretical Perspectives and Criteria for Selective Integration. *Journal of Management Accounting Research*, 15(1), 3–49.
- Datar, S. M., & Rajan, M. V. (2021). *Hornsgren's cost accounting: a managerial emphasis* (Seventeenth, global edition.). Pearson.
- Devlin, J., Ming-Wei, C., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv.Org.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm Aversion: People erroneously avoid Algorithms after seeing them Err. *Journal of Experimental Psychology: General*, 144(1), 114–126.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming Algorithm Aversion: People will use imperfect Algorithms if they can (even slightly) Modify them. *Management Science*, 64(3), 1155–1170.
- Di Vaio, A., Hassan, R., & Alavoine, C. (2022). Data intelligence and analytics: A bibliometric analysis of human–Artificial intelligence in public sector decision-making effectiveness. *Technological Forecasting & Social Change*, 174, 121201–.
- Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*.

- Dunk, A. S. (1995). The joint effects of budgetary slack and task uncertainty on subunit performance. *Accounting and Finance* (Parkville), 35(2), 61–75.
- Eisikovits, N., Johnson, W. C., & Markelevich, A. (2024). Should Accountants Be Afraid of AI? Risks and Opportunities of Incorporating Artificial Intelligence into Accounting and Auditing. *Accounting Horizons*, 1-7.
- Fahse, T. B., Blohm, I., Hruby, R., & van Giffen, B. (2022). Explanation Interfaces for Sales Forecasting. *Proceedings of the European Conference on Information Systems*, Timisoara, Romania.
- FDK-SDK. (2024). Konferenz der Kantonalen Finanzdirektorinnen und Finanzdirektoren, Budget und Rechnung der Kantone, <https://www.fdk-cdf.ch/finanzdaten> (last accessed April 17, 2024).
- Fehrenbacher, D.D., Ghio, A., & Weisner, M. (2023). Advice Utilization from Predictive Analytics Tools: The Trend is Your Friend. *European Accounting Review*, 32(3), 637-662.
- Fisher, J. G., Frederickson, J. R., & Pfeffer, S. A. (2000). Budgeting: An Experimental Investigation of the Effects of Negotiation. *The Accounting Review*, 75(1), 93–114.
- Gilad, Z., Amir, O., & Levontin, L. (2021). The Effects of Warmth and Competence Perceptions on Users' Choice of an AI System. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 583,1-13.
- He, Z., Lin, D., Lau, T., & Wu, M. (2019). Gradient Boosting Machine: A Survey. *arXiv.Org*.
- Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, J., & Baesens, B. (2011). An Empirical Evaluation of the Comprehensibility of Decision Table, Tree and Rule-based Predictive Models. *Decision Support Systems*, 51(1), 141-154.
- Jensen, M. C. (2003). Paying People to Lie: the Truth about the Budgeting Process. *European Financial Management : The Journal of the European Financial Management Association*, 9(3), 379–406.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4), 305-360.
- Liessem, T., Schedlinsky, I., Schwering, A., & Sommer, F. (2015). Budgetary Slack under budget-based Incentive Schemes—the Behavioral Impact of Social Preferences, Organizational Justice, and Moral Disengagement. *Journal of Management Control*, 26(1), 81-94.
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35-65.
- Marlowe, J. (2011). Beyond 5 Percent: Optimal Municipal Slack Resources and Credit Ratings. *Public Budgeting & Finance*, 31(4), 93-108.
- Puranam, P. (2021). Human–AI collaborative decision-making as an organization design problem. *Journal of Organization Design*, 10(2), 75-80.
- Puron-Cid, G. (2014). Factors for a successful adoption of budgetary transparency innovations: A questionnaire report of an open government initiative in Mexico. *Government Information Quarterly*, 31, S49–S62.
- Quarati, A. (2023). Open Government Data: Usage trends and metadata quality. *Journal of Information Science*, 49(4), 887–910.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in My Hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence, *Business Horizons*, 61(1), 15-25.
- Rühli, L., & Hutter, E. (2023). Budgetierung ausser Rand und Band, Avenir Suisse, <https://www.avenir-suisse.ch/publication/budgetierung-ausser-rand-und-band/> (last accessed April 19,2024).
- Secinaro, S., Calandra, D., Lanzalonga, F., & Biancone, P. (2024). The Role of Artificial Intelligence in Management Accounting: An Exploratory Case Study. In *Digital Transformation in Accounting and Auditing* (pp. 207–236). Springer International Publishing.
- Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, 61(4), 66–83.
- Sicilia, M., & Steccolini, I. (2017). Public budgeting in search for an identity: state of the art and future challenges. *Public Management Review*, 19(7), 905–910.
- Valle-Cruz, D., Fernandez-Cortez, V., & Gil-Garcia, J. R. (2022). From E-budgeting to smart budgeting: Exploring the potential of artificial intelligence in government decision-making for resource allocation. *Government Information Quarterly*, 39(2), 101644-.
- Van der Stede, W. A. (2000). The relationship between two consequences of budgetary controls: budgetary slack creation and managerial short-term orientation. *Accounting, Organizations and Society*, 25(6), 609–622.