APPLICATION OF CAUSAL MACHINE LEARNING TO COMPUTE THE EFFECTS OF VIDEO GAME INTENSITY ON STUDENTS' ACADEMIC PERFORMANCE AND PSYCHOSOCIAL WELL-BEING

Assylbek Gizatov
Senior Researcher, Research Institute of Mathematics and Applied Technologies,
Lecturer, Department of Software Engineering,
Atyrau University named after Kh. Dosmukhamedov
Atyrau, Kazakhstan
a.gizatov@asu.edu.kz; orcid.org/0009-0009-2557-0767

Nuraly Kituyev
B.A. Student, School of Sciences and Humanities
Nazarbayev University
Astana, Kazakhstan
nuraly.kituyev@nu.edu.kz; orcid.org/0009-0003-2862-8203

Izbassar Orynbassar
M.Sc. Student, Department of Software Engineering
Atyrau University named after Kh. Dosmukhamedov
Atyrau, Kazakhstan
izbassar.eng@gmail.com; orcid.org/0009-0002-4224-3249

Galymzhan Amangeldy
B.A. Student, Department of Software Engineering
Atyrau University named after Kh. Dosmukhamedov
Atyrau, Kazakhstan
gosha122804@gmail.com; orcid.org/0009-0002-7029-641X

Akerke Zhannurova
B.A. Student, School of Sciences and Humanities
Nazarbayev University
Astana, Kazakhstan
akerke.zhannurova@nu.edu.kz; orcid.org/0009-0009-9460-6778

ABSTRACT

This study applies Double/Debiased Machine Learning (DML) to estimate the quasi-causal effects of video game intensity on students' academic performance and psychosocial well-being using survey data from university students in Kazakhstan, Kyrgyzstan, and Uzbekistan. While prior research often reports negative correlations between intensive gaming and academic outcomes, such associations may reflect

Proceedings of DiGRA CA 2026

confounding and self-selection rather than causal effects. Using a causal machine learning framework that combines flexible nuisance estimation with orthogonalized treatment effect estimation, the analysis provides more credible causal inference under conditional unconfoundedness. The results indicate that the negative association between gaming intensity and academic performance is largely explained by confounding rather than a direct causal effect. However, gaming shows a consistent negative effect on sleep quality, suggesting sleep disruption as a key mechanism influencing student well-being. These findings demonstrate the importance of causal inference methods in digital games research.

Keywords

causal machine learning, double/debiased machine learning, video games, academic performance, sleep quality, stress, heterogeneous treatment effects.

INTRODUCTION

Video gaming is widespread among students and is often discussed as a potential risk factor for academic underperformance and reduced well-being. Meta-analytic evidence suggests that technology use, including video game playing, is on average negatively associated with academic performance, although the magnitude is typically small and results vary across contexts and measures (Alzahrani and Griffiths 2025; Kus 2025). Reviews further emphasise that much of the gaming–academic literature relies on cross-sectional designs, making the direction of effects difficult to establish (Hartanto, Toh, and Yang 2018). Sleep disruption is frequently discussed as a plausible pathway, with both screen exposure and gaming volume linked to poorer sleep outcomes (Exelmans and Van den Bulck 2015; Hale and Guan 2015). Gaming can also be associated with psychosocial adjustment in non-linear ways depending on intensity (Przybylski 2014).

Students differ in motivation, prior academic achievement, and life circumstances, which may simul- taneously affect both gaming behaviour and educational outcomes. These factors create substantial risks of confounding and self-selection, meaning that negative correlations may overstate direct academic harm.

Video gameplay is not, however, a monolithic behaviour. Game studies scholarship has long established that gaming encompasses highly heterogeneous practices that differ substantially by genre, platform, social context, and player motivation (Mäyrä 2008; Apperley 2006; Yee 2006). Competitive multiplayer gaming, casual mobile play, immersive single-player narratives, and social cooperative sessions involve different temporal rhythms, engagement patterns, and psychological demands. Evidence suggests that the social and contextual dimensions of play can moderate associations with well-being (Przybylski 2014), while genre-specific and platform-specific effects may diverge considerably from aggregate patterns. The present study cannot and does not speak to these distinctions: gaming intensity is operationalised solely as total weekly hours of gameplay and does not capture genre, platform, social context, or player motivation. The findings should therefore be understood as estimates of the aggregate effect of playtime volume, not as claims about any gaming practice. Applying rigorous causal inference methods to this aggregate measure addresses a significant gap in a literature dominated by cross-sectional and correlational designs and offers a methodological contribution relevant beyond psychology to the broader digital media studies community.

To address these limitations, this study uses causal machine learning. Under conditional unconfound- edness (Rosenbaum and Rubin 1983), Double/Debiased Machine Learning (DML) provides orthogonalised treatment effect esti- mates with valid inference in high-dimensional settings (Chernozhukov et al. 2018). Heterogeneous effects are explored using Causal Forest / generalized random forest ideas (Athey, Tibshirani, and Wager 2019; Wager and Athey 2018). We estimate the effects of gaming intensity on academic performance, sleep quality, and stress among university students in Kazakhstan, Kyrgyzstan, and Uzbekistan.

METHODOLOGY

The empirical strategy follows a transparent sequence to separate correlational patterns from quasi- causal effects. We (i) construct and standardise key variables, (ii) replicate conventional evidence via correlations and OLS regressions, (iii) estimate orthogonalised Average Treatment Effects (ATE) using Double/Debiased Machine Learning (DML), and (iv) examine heterogeneity via Causal Forest estimates of Conditional Average Treatment Effects (CATE).

Data Description

The empirical analysis is based on cross-sectional survey data collected via Google Forms from university students in Kazakhstan, Kyrgyzstan, and Uzbekistan. Participation was anonymous and vol- untary. The final sample consists of 2,267 observations.

Ethics statement. The survey was conducted on a voluntary and fully anonymous basis; no personally identifiable information was collected or retained. Participants were informed of the study purpose prior to completing the questionnaire and were free to withdraw at any time without consequence. All responses were aggregated for statistical analysis, and no individual-level data is reported. The study did not involve minors, medical or clinical procedures, or categories of sensitive personal data beyond standard academic and lifestyle survey items.

The dataset includes demographic variables (age, gender), academic indicators (self-reported GPA categories), gaming behaviour (hours per day and days per week), and psychosocial outcomes measured on Likert-type scales, including sleep quality and stress. Additional variables capture external constraints affecting gaming behaviour. Gaming intensity is constructed as weekly hours of gameplay. Figure 1 summarises the distribution of gaming intensity in the sample.

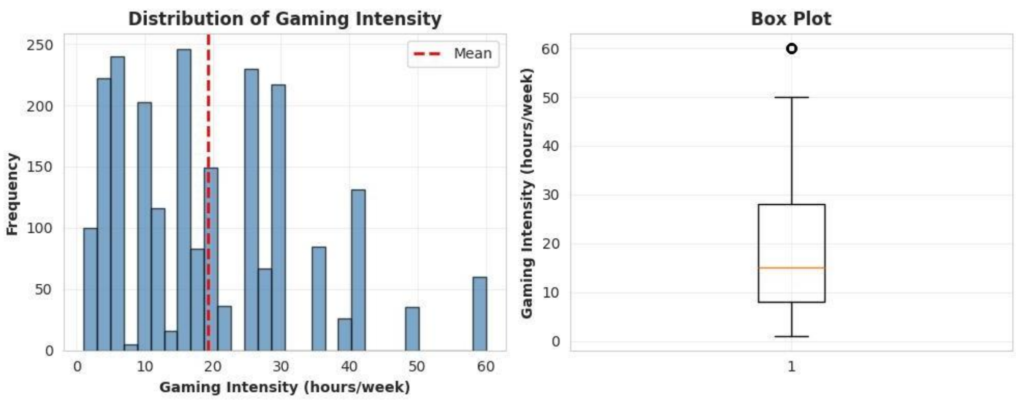

Figure 1: Gaming intensity (hours/week): histogram and box plot

Empirical strategy

Let T_i denote gaming intensity and Y_i denote an outcome. The parameter of interest is the Average Treatment Effect (ATE):

$$\theta = E[Y_i(1) - Y_i(0)]$$

Because gaming intensity is not randomly assigned, identification relies on conditional unconfoundedness:

$$(Y_i(1), Y_i(0)) \perp T_i | X_i$$

where X_i denotes observed pre-treatment covariates.

Variable construction

Gaming intensity is constructed as total weekly hours:

$$T_i = (\text{Hours per day})_i \times (\text{Days per week})_i$$

Outcomes include academic performance (GPA midpoint values), sleep quality, and stress:

$$Y_i \in \backslash\{Academic_i, Sleep_i, Stress_i\backslash\}$$

The covariate set X_i includes demographic characteristics and background variables that may influence both gaming behaviour and outcomes.

Baseline correlational analysis

We first compute pairwise correlations between gaming intensity and each outcome. These estimates describe raw associations but do not address confounding or reverse causality, motivating regression- based and causal machine learning approaches.

Ordinary Least Squares (OLS)

We estimate naive linear regressions:

$$Y_i = \alpha + \beta T_i + \varepsilon_i,$$

and covariate-adjusted specifications:

$$Y_i = \alpha + \beta T_i + \gamma X_i + \varepsilon_i.$$

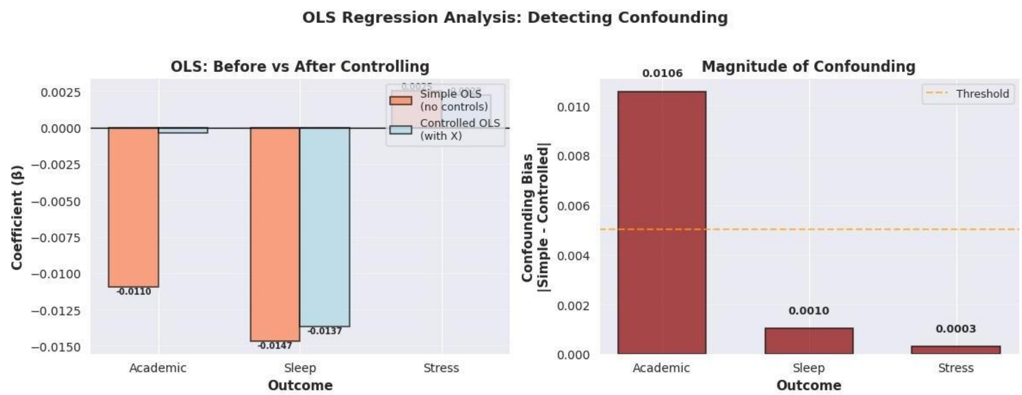

Figure 2: OLS: naive vs. controlled estimates (confounding diagnostic).

Figure 2 shows that the academic coefficient shrinks substantially after including covariates, indi- cating that naive academic "penalties" may reflect selection and confounding. By contrast, the sleep coefficient remains more stable, while the stress coefficient is small in both specifications.

Double/Debiased Machine Learning (DML)

To obtain orthogonalised causal estimates, we implement DML using the partially linear model:

$$Y_i = \theta T_i + g(X_i) + \varepsilon_i\,, \qquad T_i = m(X_i) + v_i$$

where g (·) and m (·) are nuisance functions estimated via machine learning. Residualisation yields:

$$\widetilde{Y_i} = Y_i - \hat{g}(X_i)\widetilde{T_i} = T_i - \widehat{m}(X_i)$$

and the ATE is estimated by:

$$\hat{\theta} = \frac{\sum_{i=1}^{n} \widetilde{T_i}\widetilde{Y_i}}{\sum_{i=1}^{n} \widetilde{T_i^2}}$$

Cross-fitting is used to reduce overfitting bias and preserve valid inference.

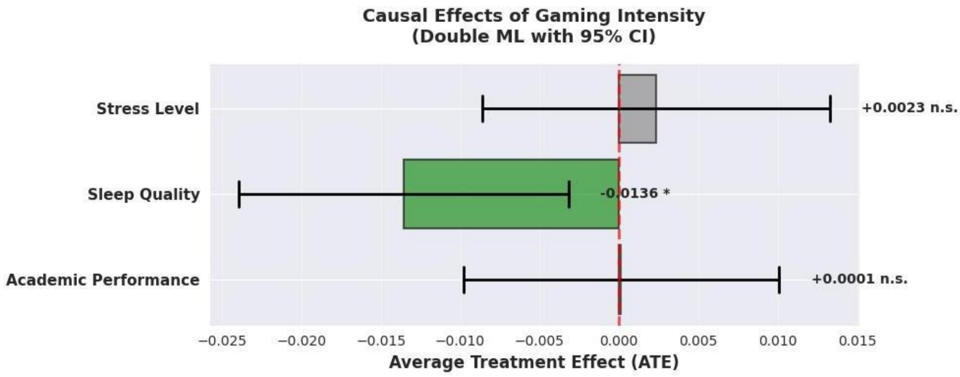

Figure 3: DML estimates of ATE with 95% confidence intervals

Heterogeneous treatment effects (causal forest)

To explore heterogeneity, we estimate conditional treatment effects:

$$\tau(x) = E[Y(1) - Y(0)|X = x],$$

using Causal Forest / generalized random forest methods.

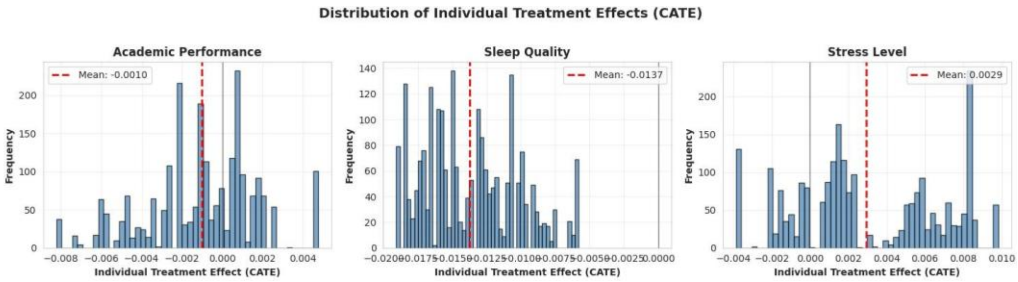

Figure 4: CATE distribution (Causal Forest).

RESULTS AND DISCUSSION

Figure 2 summarises how conventional regression evidence changes once observable covariates are included. In naive specifications, gaming intensity appears negatively related to academic performance and sleep quality, while the association with stress is weak. After adjustment, the academic coefficient moves markedly toward zero, indicating that much of the apparent "academic harm" is explained by selection and confounding rather than a stable structural relationship.

We then turn to causal machine learning. Figure 3 reports DML estimates of the ATE for each outcome. The estimated causal effect on academic performance is statistically indistinguishable from zero, implying that the negative academic association found in correlations and naive OLS is largely non-causal. In contrast, the effect on sleep quality remains negative and statistically significant under DML, identifying sleep disruption as the most robust channel through which intensive gaming affects student well-being. For stress outcomes, the estimated average causal effect is small and not statistically significant.

Finally, figure 4 shows the distribution of CATE estimates. The academic effect distribution is cen- tred near zero, reinforcing the absence of a systematic academic penalty. The sleep effect is consistently negative across individuals with limited dispersion, suggesting that heterogeneity is present but modest in practical magnitude. No clear subgroup-specific patterns emerge for stress.

Overall, the sequential comparison - from correlations to OLS to DML and CATE demonstrates that conventional regression-based evidence can overstate academic harms of gaming, while causal machine learning provides a more policy-relevant conclusion: the primary robust impact of intensive gaming is on sleep rather than academic performance.

From a game studies perspective, the robustness of sleep disruption across all estimation strategies carries direct implications for the design and governance of digital play environments. Unlike academic performance - where the confounding story redirects attention to player demographics and pre-existing study habits rather than gaming per se - sleep disruption appears to be a real downstream consequence of sustained playtime volume. This pattern is consistent with screen-based temporal displacement hypotheses, whereby late-evening gaming sessions compress available sleep time irrespective of the specific game or motivation involved (Exelmans and Van den Bulck 2015). For game designers and platform operators, the finding points toward mitigation strategies that do not require restricting play as such. Interface-level interventions - in-game session-duration reminders, evening wind-down notifications, opt-in playtime dashboards, and platform-level bedtime warning systems - could plausibly reduce sleep displacement without curtailing the experiences players seek (Shaw 2010). Such approaches align with emerging practices in player-welfare-conscious game design. It must be emphasised, however, that this study estimates the effect of aggregate weekly hours only; whether the sleep disruption mechanism operates uniformly across late-night competitive gaming, social play, or casual sessions remains an open empirical question for future research.

CONCLUSION

This study applies causal machine learning to estimate the effect of gaming intensity operationalised as total weekly hours of gameplay - on academic performance and psychosocial outcomes among university students in Kazakhstan, Kyrgyzstan, and Uzbekistan. Consistent with game studies scholarship that foregrounds the diversity of digital play (Mäyrä 2008; Apperley 2006), the analysis does not treat gaming as a monolithic activity; its conclusions are bounded to the aggregate effect of playtime volume and do not extend to genre-specific, motivation-specific, or platform-specific gaming contexts. Within this scope, correlational evidence suggests that intensive gaming is associated with lower academic performance, but causal estimates using Double/Debiased Machine Learning indicate that this association is largely an artefact of confounding and self-selection rather than a direct causal effect.

By contrast, gaming intensity shows a consistent negative causal impact on sleep quality - an effect that is stable across both average and heterogeneous treatment effect estimates. Sleep disruption thus emerges as the primary pathway through which high playtime volume affects student well-being, while no systematic average effect on stress is observed. These findings suggest that concerns about gaming and student welfare are most productively focused on the sleep disruption mechanism rather than academic performance per se. At the design and policy level, this points toward targeted interventions in-game session timers, platform-level bedtime notifications, playtime transparency tools — that address the temporal displacement of sleep without pathologising play more broadly.

Methodologically, the study demonstrates the importance of credible causal identification in digital media research. Conventional regression approaches overstate academic harm when confounding from student characteristics is not adequately controlled. Double/Debiased Machine Learning, by separating nuisance estimation from treatment effect estimation, provides a more robust and policy-relevant inference. Future research should extend this framework to genre- and context-specific play patterns, combine self-reported survey data with behavioural playtime logs, and explore whether design-level interventions can demonstrably reduce the sleep disruption mechanism identified here.

REFERENCES

Alzahrani, A. K. D., and M. D. Griffiths. 2025. “Problematic Gaming and Students’ Academic Performance: A Systematic Review.” International Journal of Mental Health and Addiction 23: 4062–4095. https://doi.org/10.1007/s11469-024-01338-5.

Apperley, Thomas H. 2006. "Genre and Game Studies: Toward a Critical Approach to Video Game Genres." *Simulation & Gaming* 37 (1): 6–23. https://doi.org/10.1177/1046878105282278.

Athey, S., J. Tibshirani, and S. Wager. 2019. “Generalized Random Forests.” The Annals of Statistics 47 (2): 1148–1178. https://doi.org/10.1214/18-AOS1709.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. 2018. “Double/Debiased Machine Learning for Treatment and Structural Parameters.” The Econometrics Journal 21 (1): C1–C68. https://doi.org/10.1111/ectj.12097.

Exelmans, L., and J. Van den Bulck. 2015. “Sleep Quality Is Negatively Related to Video Gaming Volume in Adults.” Journal of Sleep Research 24 (2): 189–196. https://doi.org/10.1111/jsr.12255.

Hale, L., and S. Guan. 2015. “Screen Time and Sleep among School-Aged Children and Adolescents: A Systematic Literature Review.” Sleep Medicine Reviews 21: 50–58. https://doi.org/10.1016/j.smrv.2014.07.007.

Hartanto, A., W. X. Toh, and H. Yang. 2018. “Context Counts: The Different Implications of Weekday and Weekend Video Gaming for Academic Performance in Mathematics, Reading, and Science.” Computers & Education 120: 51–63. https://doi.org/10.1016/j.compedu.2017.12.007.

Kus, M. 2025. “A Meta-Analysis of the Impact of Technology Related Factors on Students’ Academic Performance.” Frontiers in Psychology 16: 1524645. https://doi.org/10.3389/fpsyg.2025.1524645.

Mäyrä, Frans. 2008. *An Introduction to Game Studies: Games in Culture*. London: SAGE.

Przybylski, A. K. 2014. “Electronic Gaming and Psychosocial Adjustment.” Pediatrics 134 (3): e716–e722. https://doi.org/10.1542/peds.2013-4021.

Rosenbaum, P. R., and D. B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” Biometrika 70 (1): 41–55. https://doi.org/10.1093/biomet/70.1.41.

Shaw, Adrienne. 2010. "What Is Video Game Culture? Cultural Studies and Game Studies." *Games and Culture* 5 (4): 403–424. https://doi.org/10.1177/1555412009360414.

Wager, S., and S. Athey. 2018. “Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests.” Journal of the American Statistical Association 113 (523): 1228–1242. https://doi.org/10.1080/01621459.2017.1319839.

Yee, Nick. 2006. "Motivations for Play in Online Games." *CyberPsychology & Behavior* 9 (6): 772–775. https://doi.org/10.1089/cpb.2006.9.772.