Efficiency in Virtual Markets:

Validating Assumptions of Information in OldSchool RuneScape

David Lundgren (lundgre1)

Daniel Huang (huangd22)

Alexander Tran (tranal11)

ECO481H1: Special Topics in Economics with Data Analytics

Professor: Marlène Koffi

10 pages (excluding citations)

1. Introduction

Academic interest surrounding artificial economies has grown significantly as economists acquiesce in virtual markets' ability to model and reflect traditional market behaviour (Bainbridge, 2007; Bartle, 2006; Ondrejka, 2006). Agent-based computational simulations and virtual models of inflation, shocks, and policy have utilised this convergence to extrapolate findings to real-world situations (Humphreys, 2008). Indeed, as quantitative economic analysis grows prominent in cross-disciplinary fields, significant attention should be devoted towards evaluating the economic assumptions that underlie theory in virtual laboratories (Lehdonvirta, 2010). We find that there exists a scarce amount of literature that tests the assumptions that characterise virtual markets. This vacuum could significantly hinder the performative power of virtual laboratories as unscrutinised economic assumptions limit the applicability of findings to reflect real-world dynamics.

We examine the persistence and dissipation of information in virtual economies – mainly, whether the assumption of perfect information holds in virtual settings. The Efficient Market Hypothesis (EMH) postulates that prices in a market reflect all available information, and that any short-term changes in price follow a 'walk' dominated by random innovations (Hewamalage et al., 2022; Rossi, 2005). Research widely suggests that online, agricultural, foreign exchange, and other traditional markets are subject to the EMH (Fama, 1970; French & Fama, 2007, Lo & MacKinlay, 1988; Malkiel, 2003;). As such, determining whether the theory holds in virtual settings can inform economists about the feasibility of virtual markets to model real-world economic phenomena, and further inform policymakers about the effects of economic shocks as informational releases.

This research looks at the virtual world of Old School RuneScape (OSRS), a large online multiplayer game with a virtual economy that adequately converges to real-world market behaviour and dynamics (Heeks, 2010). The game boasts a large fanbase of 39 million users, 2.9

million of which daily actively engage in the production, trade, and consumption of virtual goods for gameplay purposes (*Server Population & Player Count,* 2023). The estimated trading volume is over £60 million in real value as the virtual player-to-player market makes up an integral and necessary part of the game (Scholten et al., 2019). Specifically, this virtual world is of particular interest to our analysis because of a new game update that is hypothesised to alter market forces (Gonzalez, 2023). The update, which literature suggests can assimilate to a regulatory intervention or exogenous shock, is hypothesised to increase the demand for and price of dragon bones (Hogan-Hennessy et al., 2022). Therefore, by analysing how the price of dragon bones changes post-update in simulation of an economic shock, we can examine how the virtual market responds to the injection of new information (EMH; Yang, 2013). As an extension, this research also examines whether it is possible to predict the direction of price change. This analysis could draw similarities (or dissimilarities) between functional investment strategies that utilise informational asymmetries for profiteering in traditional markets, and whether that is possible in virtual markets. Therefore, in summary, our research is three-fold: (1) has the game-update been received by the player community; (2) does the efficient market hypothesis hold; and, (3) can we predict price?

The framework above linearizes our study on whether the assumption of perfect information holds in virtual markets. By placing under scrutiny the assumptions that economic theory and associated experiments rest upon, we indirectly lend credence to how scholars should approach virtual markets in the future and to what extent results can be extrapolated to the real world. The first question concerns a broader commentary on how market interventions are perceived by economic agents and lays the foundation for whether the game-update *should* have had an impact on prices. If we affirm that the update (shock) has impacted players (economic agents), then we can study how the market responds to the injection of new information. This analysis is captured in questions (2) and (3), where we predict and test whether prices follow a random walk to determine if behaviour is dominated by random shocks. We foresee that the hypothesis holds and that prices follow a random walk, since informational asymmetries are likely instantaneously corrected by the digital market forces that associate price (Yang et al., 2021).

Relevant literature focuses on virtual markets to document and examine real-world behaviour. Specifically, there is research that relies upon OSRS to model traditional market dynamics. Scholten et al. (2019) constructs price indexes for virtual goods in OSRS to study virtual economies, inflation, and stability. The results show that OSRS has a healthy economy that converges with real economic trends. Further, the authors use web-scraping to obtain OSRS price data– a method which is replicated in this paper. Bilir (2009) compares virtual economies to real markets, focusing particularly on the assumptions that underlie virtual worlds. Her findings show that virtual economies follow a faucet-drain model and can be characterised by perfect competition with equal access, homogenous products, price-takers, and perfect information; an evaluation of the latter is precisely what this paper examines. Bilir also details literature on the capacity of virtual economies to model economic shocks and regulatory changes through game-updates. Hogan-Hennessey et al. (2022) expands upon this by relating game-updates to real-world policy changes, showing the effect of market interventions on economic shocks. The authors use causal inference to study transaction taxes, price floors, and item sinks in OSRS, finding that inflationary pressures lead to no change in trading volume. The results align with economic theory to suggest that virtual markets can accurately model policy changes that affect real economies.

However, while these papers support the use of OSRS to critically evaluate the EMH assumption, they are broader in scope, provide no commentary on the EMH, and do not use any ML techniques. This is the vacuum of research alluded to earlier which our paper attempts to fill.

2. Data and Methodology

The data in our analysis includes text and time-series data. The text data is a corpus of comments from Reddit's official discussion thread relating to the game update (*Ruinous Powers,* 2023). We collect around 2,200 comments which are then classified as documents. To prepare the text data for sentiment analysis, we first pre-process the comments. This process includes removing common stop words and punctuations provided in the NLTK package. We also remove abbreviations of certain phrases and internet slang that commonly appear in online forums (such as TBH, IMO, etc.). Once sufficiently clean, we convert the pre-processed documents into a bag of words and bigrams. We foresee that the sentiment analysis component could be affected by bigrams of in-game content, which may inaccurately characterise our sentiment scores. However, these terms are included in our analysis to contextualise the topic selection of our LDA model.

The time-series data gathers the total trade volume and average price of a set of goods affected by the game-update in 5-minute intervals (*OSRS Wiki*, 2023). Research suggests that dragon bones were most affected as the update specifically changes their effect. Additionally, the most common substitutes for dragon bones were dagannoth-, wyvern-, hydra-, and wyrm bones. Since players use these substitutes for several reasons, such as cheaper alternatives or offering more experience points (XP), we include these in our analysis. Other substitute goods are not listed as they offer a lower experience rate or were prohibitively more expensive in large quantities.

The time-series consists of 19 days captured in 5-minute intervals, with a total of 5,300 observations. While exploring the data, we observe some notable trends. The average trade volume of dragon bones far exceeds its closest substitutes. Dragon bones traded, on average, 9,000 units every 5 minutes, as opposed to several hundred units of the closest substitute. While these goods function as substitutes, our exploratory data analysis suggests that focusing our analysis on dragon bones is correct in order to analyse broader market trends in a virtual setting.

To conduct the sentiment analysis, we construct an LDA model and find the optimal number of topics using Genism. We tune the hyperparameter representing number of topics by selecting the topic with the highest coherence score based on sudden troughs or consistent drop-offs (Gan & Qi, 2021). We assign the dominant topic score to each document. Finally, we conduct our sentiment analysis on the individual documents before aggregating the sentiment scores by topic. If the individual topics all have significant positive compound scores (> 0.05), we propose that the player reaction is positive overall, and vice-versa if negative scores are below –0.05.

Once we execute our initial sentiment analysis and answer whether prices *should* change, we proceed to determine whether the EMH holds in regression and classification exercises. To evaluate this, we design a method to test for random walks. This approach has been used previously to determine efficiency in traditional markets (Bacchetta et al., 2009; Hewamalage et al., 2022; Rossi, 2005). Notably, we attempt to predict continuous prices in the regression case and direction of price change in the classification model – both utilizing nowcasting. For each, we compare the ML models to the persistence model ($Y_{t+1} = Y_t + \varepsilon_t$). If the ML models perform worse than the persistence model, then there is random walk in the price sequence. The predictor variables are the 5-minute lagged trade volumes and lagged prices for our main and substitute goods.

For the regression component, we incorporate decision tree, random forest, gradient boosted tree, and KNN models on an 80/20 training-test split. We compare the test MSE for the ML models to the test MSE of the persistence model. If the persistence model performs best by having the lowest test MSE, we conclude that there is a random walk in prices. For the classification exercise, the outcome variable is encoded to indicate the direction of price change in the current period, with 1 indicating an increase and 0 indicating decrease or stationary. The predictor variables are similar, except in the case of lagged prices which is converted in similar categorical fashion. In terms of models, we construct a decision tree, random forest, gradient boosted tree, KNN and logistic regression models. We compare the accuracy scores and AUCs of our ML models to the persistence model. Again, if these measures are higher for our persistence model, we conclude that there is random walk, and the efficient market hypothesis holds.

3. Results

The results of the sentiment analysis indicate that the official update serves as an information injection for the market to process. If the overall topical reactions are non-neutral, then we hypothesise that prices should be non-stationary in the presence of EMH. This means that our 'economic shock', the game-update, should have impacted market participants enough to affect a change in prices. First, the results of tuning our LDA model show that the number of optimal topics is 4 – as seen in Figure 1. While the coherence score returns to a comparable level (0.39) in topics 7 and 12, we select the minimum optimal number of topics to reduce model complexity (Gan & Qi, 2021). Thereafter, using contextual research, we define the four different topics as shown in Figure 2 and their associated word distributions. Table 1 shows the sentiments towards each topic.

The results show that each topic was, on average, received positively by OSRS' community. It is notable that Topic 1 represents 'Gameplay Downsides' to the new update, which suggests that a positive score of 0.22 is actually indicative of a negative sentiment towards the update. This is somewhat expected as the update enforces major changes to the game design that reforms conventional norms and player experience. Therefore, it is likely that a considerable amount of our population of interest is unhappy or reluctant to welcome changes to the game. This relates to a bearish market sentiment in real-world settings as informational releases such as regulatory or policy initiatives often result in negative receptions. Topics 2 and 3 were received relatively positively by OSRS players, which suggests a split in attitudes across the OSRS player base. Topic 4 borders neutral-territory as it is within the [-0.05, 0.05] sentiment score region.

Having determined that the new information (game-update) was received by the player base, we turn to examine how prices reflect the new information. Specifically, does the EMH hold in continuous price cases? To evaluate this query, we fit a selection of ML models on our training dataset and compare the test MSE with the performance of the persistence model. The results are shown in Table 2. As is evident, the persistence model performs significantly better, with a test MSE of 928.5 which is 23% lower than the second-best performing model. If a naïve forecast, such as persistence, represents the best prediction of price, then that series follows a random walk and is efficient (Hewamalage et al. 2023). The results therefore suggests that the EMH does hold.

The persistence model is given by: $Y_{t+1} = Y_t + \varepsilon_t$, which implies that any changes in price are a consequence of random innovations from one period to the next. Therefore, while price sequences in our case are not entirely random, fluctuations in prices are. Therefore, the price of dragon bones in OSRS does appear to follow a random walk since the best predictive model is one that presumes that variability in price from one period to the next is a product of random shocks. Research supports our methodology that this thereafter implies that the EMH does hold (Rossi, 2005: 61). The implication of this finding is that the assumption of perfect information seemingly holds in large virtual markets. This is important as it credits virtual markets with the capacity to model economic theories and reflect computational simulations of economic shocks that can encapsulate real-world trends. Therefore, buttressing the validity of this assumption allows researchers to rely on findings from virtual worlds to extrapolate onto weakly efficient conventional markets, such as agricultural, online, and foreign exchange markets.

Finally, the results of the classification component are highlighted in Table 3. Notably, categorical prices do not follow a random walk. Ergo, the efficient market hypothesis does not hold. This is because the persistence model performs much worse than any of the ML models, with an accuracy score and AUC of 0.38 – significantly worse than random. This finding contrasts the results of our continuous-price analysis but can be understood intuitively. The persistence model performs poorly in categorical cases because the variation in price is discrete. This renders the model inherently less viable to exhibit random innovations given the binary nature of the outcome - whereas continuous variables are inherently more random by definition. While this may question the validity of our efficiency finding in a categorical context, it shifts our focus to the predictive power of ML models to forecast direction of price changes following shocks, i.e., the game-update.

The best model appears to be the logistic regression model, with an accuracy score of 0.607 and AUC of 0.649. It marginally outperforms the gradient boosted tree, with scores of 0.601 and 0.636, respectively. The confusion matrix of the logistic model is depicted in Figure 3. The results show that our model performs significantly better than random chance, with a precision rate of 0.63. The recall rate (0.57) is notably worse than the true negative rate (0.64), which suggests that our model fails to capture true instances of price increases. The model predicts 182 false positives, however, predicts 234 false negative instances. This suggests that, while performing better than random, our model fails to capture the increased significance of dragon bones that was ushered in as part of the new update. Dragon bones became an approximate 'necessity' to levelling up, which our model does not account for since it is only trained on market-level data and not sentiment-level data. In short, the model is therefore underestimating the positive effects of the economic shock on price. However, the logistic model still manages to predict the direction of price change which supports that the efficient market hypothesis does not hold in the categorical case.

4. Conclusion

This research confirms that there is evidence to support the economic assumptions that premise studying virtual markets as a proxy for real-world markets. This paper utilizes ML methods to test the validity of the EMH in the context of OSRS' virtual market following an economic shock. Sentiment analysis shows that prices do indeed adjust to the information release, and nowcasting is effective at predicting the direction of these changes better than random chance. The results suggest that EMH holds in continuous virtual settings, which is consistent with expectations and previous research. It is worth noting that we do not investigate the existence of behavioural biases or informational asymmetries, given the difficulties involved in measuring these effects. However, both are fundamental discrepancies to our assumptions of virtual and real-world market spaces. Further work targeting these violations would be a valuable contribution for future research.

This study confirms the robustness of foundational assumptions in using virtual markets as laboratories for economic analysis, which serves to alleviate scholarly scepticism surrounding its validity. In addition, we observe that OSRS' market behaves similar to our expectations of real-world markets: agents are utility-oriented, attempt to maximize their utility, and adjust expectations and react accordingly in the event of a shock. Further, virtual markets are facilitated by entirely digital information systems, which reflects a growing trend towards digitalization.

To conclude, we find that OSRS' virtual market does not violate the EMH, and supports the premise that virtual markets can model real-world economic outcomes. However, we caution that economists should not use virtual settings for economic experiments without scrutiny, and should diligently confirm these assumptions to extrapolate their findings to conventional contexts. All said, given the progressive digitalization of global economies and virtual marketplaces, we expect virtual settings to become an increasingly valuable asset in future economists' toolbox.

5. Appendix

Figure 1: LDA Coherence Scores

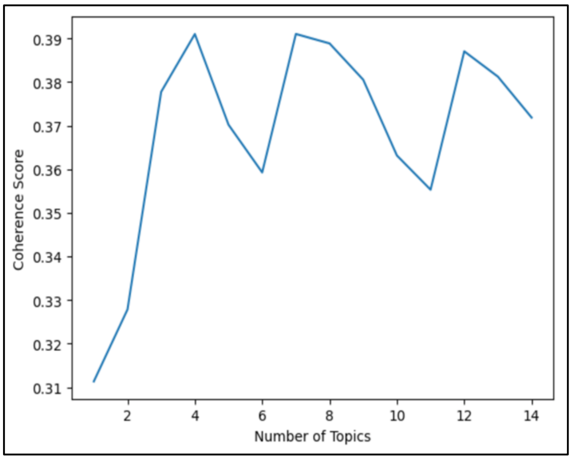


Figure 2: LDA Topics

Table 1: Sentiment Analysis Results

	Negative Score	Neutral Score	Positive Score	Compound Score
Gameplay Downsides	0.125570	0.640641	0.215581	0.221394
Defensive Applications	0.134134	0.697675	0.168188	0.061254
Gameplay Changes	0.154706	0.646766	0.198525	0.105041
Combat Application	0.162159	0.642311	0.195533	0.049417

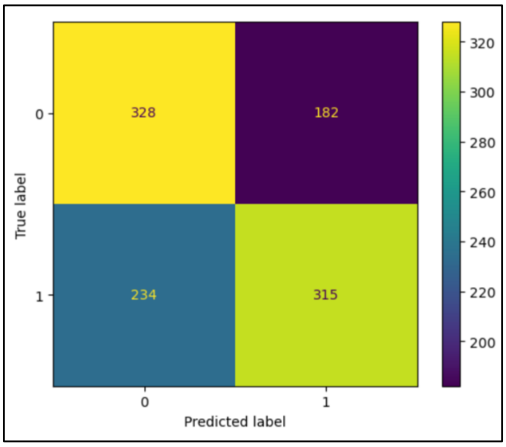
Table 2: Random-walk Test for Continuous Prices

Model	Test MSE
Persistence Model	928.500905
Decision Tree Model	1655.198362
Random Forest Model	1204.756096
Gradient Boosted Tree	1315.572298
KNN	1478.253437

Table 3: Random-walk Test for Categorical Prices

Model	Accuracy Score	AUC
Logistic Regression	0.607177	0.639121
Decision Tree	0.543909	0.544091
Random Forest	0.583569	0.611511
Gradient Boosted Tree	0.601511	0.636830
Persistent Model	0.386213	0.385380
KNN	0.578848	0.597109

Figure 3: Confusion Matrix for Logistic Regression



6. Citations List

Bacchetta, Philippe, Elmar Mertens, and Eric van Wincoop. 2009. "Predictability in Financial Markets: What Do Survey Expectations Tell Us?" *Journal of International Money and Finance* 28 (3): 406–26. doi:10.1016/j.jimonfin.2008.09.001.

Bainbridge, William Sims. 2007. "The Scientific Research Potential of Virtual Worlds." *Science* 317 (5837): 472–76. doi:10.1126/science.1146930.

Bartle, Richard A. 2006. *Designing Virtual Worlds*. Berkeley, CA: New Riders.

Bilir, Tanla E. 2009. "Real Economics in Virtual Worlds: A Massively Multiplayer Online Game Case Study: Runescape." *SSRN Electronic Journal*. doi:10.2139/ssrn.1655084.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25 (2): 383. doi:10.2307/2325486.

Fama, Eugene F., and Kenneth R. French. 2007. "Dissecting Anomalies." *SSRN Electronic Journal*. doi:10.2139/ssrn.911960.

"Food." 2023. *Old School RuneScape Wiki*. Accessed April 12. https://oldschool.runescape.wiki/w/Food. Referred to as "*OSRS Wiki*."

Gan, Jingxian, and Yong Qi. 2021. "Selection of the Optimal Number of Topics for LDA Topic Model—Taking Patent Policy Analysis as an Example." *Entropy* 23 (10): 1301. doi:10.3390/e23101301.

Gonzalez, Christina. 2023. "Old School Runescape's Proposed First-Ever New Prayer Book Gets Robust, but Divided, Poll Response." *MMORPG.com*. March 3. https://www.mmorpg.com/news/old-school-runescapes-proposed-first-ever-new-prayer-book-gets-robust-but-divided-poll-response-2000127433.

Heeks, Richard. 2010. "Understanding 'Gold Farming' and Real-Money Trading as the Intersection of Real and Virtual Economies." *Journal of Virtual Worlds Research* 2 (4).

Hewamalage, Hansika, Klaus Ackermann, and Christoph Bergmeir. 2022. "Forecast Evaluation for Data Scientists: Common Pitfalls and Best Practices." *Data Mining and Knowledge Discovery* 37 (2): 788–832. doi:10.1007/s10618-022-00894-5.

Hogan-Hennessy, S., P. Xenopoulos, and C. Silva. 2022. "Market Interventions in a Large-Scale Virtual Economy." *Cornell University Arxiv*, October. doi: arXiv:2210.07970 .

Humphreys, Sal. 2008. "Ruling the Virtual World: Governance in Massively Multiplayer Online Games." *European Journal of Cultural Studies* 11 (2): 149–71. doi:10.1177/1367549407088329.

Lehdonvirta, Vili. 2010. "Virtual Economics: Applying Economics to the Study of Game Worlds." *Proceedings of the 2005 Conference on Future Play*, June. doi:https://ssrn.com/abstract=1630302.

Lo, Andrew W., and A. Craig MacKinlay. 1988. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." *The Review of Financial Studies* 1 (1): 41–66. doi:https://www.jstor.org/stable/2962126.

Malkiel, Burton G. 2003. "The Efficient Market Hypothesis and Its Critics." *Journal of Economic Perspectives* 17 (1): 59–82. doi:10.1257/089533003321164958.

Ondrejka, Cory. 2006. "Finding Common Ground in New Worlds." *Games and Culture* 1 (1): 111–15. doi:10.1177/1555412005281915.

Rossi, Barbara. 2005. "Testing Long-Horizon Predictive Ability with High Persistence, and the Meese-Rogoff Puzzle." *International Economic Review* 46 (1): 61–92. doi:https://www.jstor.org/stable/3663588.

"Ruinous Powers V1.1 - Updated Greenlight Poll Blog." 2023. *Reddit*. March. https://www.reddit.com/r/2007scape/comments/11ltva6/ruinous_powers_v11_updated_greenlight_poll_blog/. Referred to as "*Ruinous Powers.*"

Scholten, Oliver James, Peter Cowling, Kenneth A. Hawick, and James Alfred Walker. 2019. "Unconventional Exchange: Methods for Statistical Analysis of Virtual Goods." *2019 IEEE Conference on Games (CoG)*. doi:10.1109/cig.2019.8847969.

"Server Population & Player Count." 2023. *MMO Populations*. April. https://mmo-population.com/r/2007scape.

Yang, Linchuan, K.W. Chau, and Yang Chen. 2021. "Impacts of Information Asymmetry and Policy Shock on Rental and Vacancy Dynamics in Retail Property Markets." *Habitat International* 111: 102359. doi:10.1016/j.habitatint.2021.102359.

Yang, Ruoke. 2013. "Could the Virtual Be Similar to the Real? A First Look from an Efficient Markets Perspective." *Quarterly Journal of Finance* 03 (03n04): 1350019. doi:10.1142/s2010139213500195.