

Penny Auctions are Predictable

Predicting and Profiling User Behavior on DealDash

[†]Xinyi Zhang, [§]Shawn Shan, [†]Shiliang Tang, [§]Haitao Zheng, [§]Ben Y. Zhao
{xyzhang, shiliang_tang}@cs.ucsb.edu, {shansixioing, htzheng, ravenben}@cs.uchicago.edu
[†]UC Santa Barbara, [§]University of Chicago

ABSTRACT

We study user behavior and the predictability of penny auctions, auction sites often criticized for misrepresenting themselves as low-price auction marketplaces. Using a 166-day trace of 134,568 auctions involving 174 million bids on DealDash, the largest penny auction site in service, we show that a) both the timing and source of bids are highly predictable, and b) users are easily classified into clear behavioral groups by their bidding behavior, and such behaviors correlate highly with the eventual profitability of their bidding strategies. This suggests that penny auction sites are vulnerable to modeling and adversarial attacks.

CCS CONCEPTS

• **Applied computing** → **Online auctions**; • **Human-centered computing** → *User models*;

KEYWORDS

online auctions; user behavior; sequence prediction; clustering

ACM Reference Format:

[†]Xinyi Zhang, [§]Shawn Shan, [†]Shiliang Tang, [§]Haitao Zheng, [§]Ben Y. Zhao. 2018. Penny Auctions are Predictable: Predicting and Profiling User Behavior on DealDash. In *HT '18: 29th ACM Conference on Hypertext and Social Media, July 9–12, 2018, Baltimore, MD, USA*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3209542.3209576>

1 INTRODUCTION

Penny auctions, also known as pay-per-bid auctions, are auctions in which all participants must pay a bidding fee each time they place an incremental bid. Each auction starts with a small reserve price, and a countdown timer. Each new bid increments the current price by a small fixed amount and resets the timer. When the timer expires, the participant who placed the last bid wins the auction and purchases the item with the final auction price. Penny auctions are often criticized for their misleading advertisement [9] where they use an auction's final price as the cost of the auctioned item, when, in fact, the majority of the revenue comes from the prices paid for each bid.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

HT '18, July 9–12, 2018, Baltimore, MD, USA

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5427-1/18/07...\$15.00

<https://doi.org/10.1145/3209542.3209576>

The unique structure of penny auctions is designed to generate revenue from all users, both winners and losers. They are seen as exploiting human psychological tendencies such as risk-seeking behavior [18] and the sunk-cost fallacy [3]. However, the mechanisms used, i.e. short bid timer and small incremental bids, impose specific constraints on the auction process itself. A natural question arises, do these mechanisms encourage the formation of fundamental processes driving bidding behavior? And if so, do such behaviors produce predictable patterns that can be modeled and predicted using modern machine learning tools? Success could signal the possibility of designing adversarial bidding algorithms that consistently win auctions with minimal cost of bids.

Our work tries to answer these initial questions, by studying empirical traces of bids in auctions on DealDash, the largest and one of the oldest penny auctions online today. We record details of bids and outcomes for all DealDash auctions over a 166-day period, totaling 134,568 auctions with 174 million bids from 101,936 unique users.

Using analysis of this trace, our goal is to answer two key questions. First, we want to understand what types of patterns exist in bidding behavior in DealDash auctions, and if these patterns make the overall bidding process predictable using modern machine learning models. Second, we want to understand what, if any, common strategies exist in bidding behavior, and how such strategies fare in quantifiable terms. We want to identify the most successful strategies, as a first step towards developing adversarial bidding algorithms for penny auctions.

Our analysis and experiments produce some surprising results. First, our analysis shows that most users optimize their bids in accordance with the pay-per-bid auction structure. Nearly all bids come at the last possible second before timer expiration. Users tend to exhibit repetitive bidding behavior across auctions, making sequences of bids in each auction highly predictable. Second, we use similarity analysis to cluster bidders by their bidding behavior, and identify five key categories, defined by the dominant bidding strategy most commonly observed in their bidders. Mapping these bidding behaviors to auction outcomes shows that, unsurprisingly, aggressive and persistent bidders win a disproportionately high number of auctions, and earn significant gains per auction. In contrast, low-activity bidders or those limited by budget win fewer auctions per user, and generally have trouble recouping their losses from paying for bids, resulting in net loss.

We believe our work is the first to empirically study the predictability of bidding behavior in penny auctions. Our results suggest that these systems can be gamed by adversarial strategies using bid predictions from machine learning prediction models.

2 BACKGROUND & RELATED WORK

DealDash. DealDash [1] is one of the largest and longest-running penny auction websites (since 2009). Its functionality is typical of other penny auctions. On DealDash, the typical bid fee is 12-15 cents, and the fixed price increment is restricted to 1 cent. The countdown clock expires in 10 seconds. DealDash supports a buy-it-now function, by which a losing bidder can purchase the item with a posted retail price and have all the previously placed bids refunded. DealDash also provides a helper function called BidBuddy, which is a script that takes a fixed amount of budget from the participant and automatically place bids whenever the countdown clock reaches 1 second and someone else holds the last bid.

Related Work. Prior work noted that penny auctions are very profitable for the sellers [3, 18]. Most studies of penny auctions aim at finding theoretical explanations of the high revenue, *e.g.*, information asymmetry [8], risk-loving nature of bidders [18] and sunk cost fallacy [3]. Other studies have made great effort to predict the final price of an auction, either statically [11, 21] or dynamically [13, 22], using machine learning and various modeling tools.

While significant progress has been made on developing economic models on auction process and auctioneer profits, few have examined individual bidder behavior in these systems. In contrast, our work predicts whether a bid will appear and who will place the bid, which is more similar to individual behavior models for sequence synthesis done on small-scale English auctions [6, 17].

Some studies have identified different types bidding behavior in “traditional” auction settings such as English or Dutch auctions. The behaviors identified include jumping [10], snipping [4, 19], evaluating [5], participatory [5], strategic exiting [2] and shilling [20]. To the best of our knowledge, our work is the first to systematically identify bidding behavior used in penny auction settings.

3 INITIAL ANALYSIS

We begin by describing our data collection methodology and our dataset. Then we perform preliminary analysis to understand the bidding activities on DealDash. This provides the context for in-depth studies in later sections.

3.1 Data Collection

We collect all observed auctions from DealDash over a 166-day period, from October 19th, 2017 to April 3rd, 2018. In total, we obtained complete history for 134,568 auctions. Each auction contains the name of the item, the buy-it-now price, the starting time and each bid placed during the auction. For each bid, we record the ID of the user placing this bid, and the time of the bid. We extracted 101,936 unique users from a total of 174,076,943 bids.

To verify the completeness of our dataset, we launched two sets of crawlers from two distinct sets of IP addresses. We find that 99.7% of all observed auctions are identified by both crawlers. It is thus safe to conclude that we have covered most or all of the auctions that a typical DealDash user would see.

3.2 Preliminary Analysis

For an initial understanding of the user behaviors in penny auctions, we start by answering a few basic questions.

When are bids placed? In DealDash, the countdown timer expires in 10 seconds. When looking at how long users waited before placing their bids, Figure 1 shows that the majority of bids (81.6%) are placed at the very last second. The reason of this is twofold. First, it is in the users’ best interest to wait until the last seconds in hope of someone else being impatient enough to do the bidding for them. Second, DealDash provides a functionality called *BidBuddy* which automatically place bids for the user at the last second.

How many bids do you need to win? We look at the number of bids placed by users in an auction, as well as the winning users. As shown in Figure 2, half of the users place more than an average of 20 bids per auction. To win an auction, more bids need to be placed. Only 8.1% of the users can win an auction with less than 10 bids. In half of the auctions, more than 21.3% of all bids are placed by the auction winner. This shows that winning an auction is more than being at the right place at the right time. It takes repeated bids to convince the other bidders to give up. In Section 5, we delve into detail about the behavior patterns of auction winners.

How much does an auction winner gain? As discussed in Section 1, the actual savings from penny auction are often unclear to users. We examine how much money is “won” in each DealDash auction. When calculating the value of the item won, we estimate the retail price using the price of the same item offered on Amazon.com. Note that the price of each bid ranges from 12 to 15 cents. Here we use 12 cents as an estimation of the bid cost for each user. Using 13, 14 or 15 cents per auction produces similar results.

Figure 3 shows a CDF of the proportion of money paid by the winners and all participants during each auction. In the majority of the cases, even after accounting for the price of bids, the winner pays significantly less than retail for the item won. Half of the winners win the item after spending 12.3% of the retail price. Even accounting for bids placed by the losing bidders, DealDash is only able to generate profit out of 32.5% of the auctions. One thing to note is that certain bidders will pay more than the retail price to win auctions. For example, a user called “leilani2” placed 169,223 bids on a Lawn Mower with retail price of \$3,099. The bids would cost approximately \$20,307. This behavior is commonly associated with power bidders [15]. Power bidders establish a reputation by not giving up an auction even when the price goes unreasonably high. They hope that experienced users will learn to avoid the power bidders, and they can then win auctions at a very low price. We show in Section 5 that in the case of DealDash, most power bidders still need to put in a large number of bids to win auctions.

4 BID PREDICTION

In this section we study the predictability of bidders in penny auctions. We achieve this by building machine learning models to predict bidding behavior and evaluating them using our DealDash traces. More specifically, we use n-gram [7] and recurrent neural network (2-layer LSTM [12]) models, and show that bidding behavior in penny auctions are actually highly predictable.

4.1 Methodology

Our prediction models use a sequence-based framework. An auction with n bids forms a sequence $S = (U_1, U_2, \dots, U_n, End)$, where

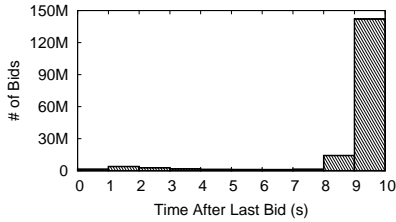


Figure 1: Distribution of time waited until the next bid.

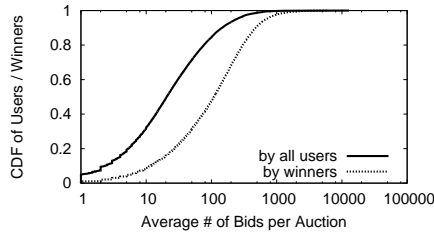


Figure 2: Bids placed per auction, for auctions winners or across all bidders.

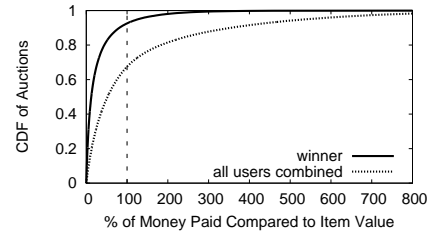


Figure 3: Distribution of price paid relative to product value.

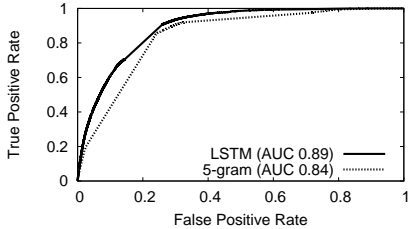


Figure 4: ROC-curve for predicting whether an auction will end.

	AUC	Top-1 Accuracy	Perplexity
3-gram	0.831	0.875	1.557
4-gram	0.843	0.886	1.653
5-gram	0.844	0.890	1.827
6-gram	0.839	0.891	2.063
7-gram	0.831	0.890	2.318
LSTM	0.890	0.900	1.259

Table 1: AUC in predicting auction ending, Top-1 accuracy and perplexity in predicting next bidder.

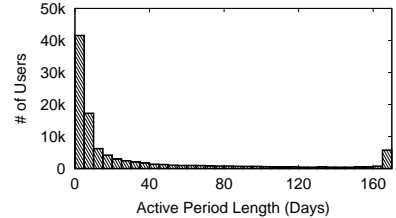


Figure 5: Distribution of time gap between each users' first and last bid.

U_i is the username of the user who placed the i^{th} bid in the auction, and End is a token that indicates auction has ended. Our model takes in any subsequence (U_1, U_2, \dots, U_k) where $k \leq n$ and predicts the value of S_{k+1} , either the username of the $k+1^{th}$ bidder or that the auction has ended.

A bidder who has recently placed a bid in an auction is likely to bid again soon. Hence, we introduce the concept of "relative position," as the position difference between the current bid and the most recent bid placed by the same user. We can then transform a bidding sequence into a sequence of relative positions of the bidder who placed each bid. When a user has not previously placed any bid in the auction, her relative position will be regarded as infinite.

For example, a bidding sequence $S = (A, B, A, C, B, A, End)$ can be transformed into $S_{rel} = (Inf, Inf, 2, Inf, 3, 3, End)$, which is then used in prediction. We find the performance is better when we treat different relative positions as separate classes instead of as a numeric value. Given that most relative bids (95.5%) are below 20, we enforce a cap of 20 on the maximum relative position. Experiments show different threshold values of 10, 20 and 50 have minimal impact on prediction results.

4.2 Prediction Results

We separate our data trace by time, auctions that concluded at least 14 days before the end of the trace (120,671 auctions), and auctions taking place during the last 14 days (13,842 auctions). We use the older auctions for training, and the newer auctions for testing.

We first use the probability on the class "End" to predict whether an auction will end or not. In Table 1, the best performing n-gram model is 5-gram with an AUC of 0.844. The LSTM model beats it by a large margin with an AUC of 0.890, as shown in Figure 4.

We evaluate bidder prediction using top-1 accuracy, which is the percent of time our model correctly predicts whom the next bidder is. As shown in Table 1, the LSTM achieves a slightly higher top-1 accuracy of 0.900. In addition, we evaluated each model's

perplexity. It is formulated as $\exp(-\sum_i (\log p_i)/n)$, where p_i is the predicted probability of the i -th relative position in the test set of n bids. A lower perplexity indicates a better model. LSTM significantly outperforms the rest.

The high accuracy in bid prediction indicates strong underlying processes that drive bidder behavior. Next, we will try to identify these underlying behavior patterns, and show how these behaviors lead to the failure or success of a penny auction bidder.

5 PROFILING USER BIDDING BEHAVIOR

To identify behavior patterns underlying how users bid, we cluster users based on similarity in statistical features of their bidding history, and then analyze the bidding performance of each cluster.

5.1 Clustering Methodology

Users' bidding patterns manifest only when there is sufficient bidding history. With 19.3% of users joining only one auction, user activity is highly skewed. As shown in Figure 5, users fall into two distinct types. The ones who played for a short period and quit, and those who are active for most of the crawling period. This high user churn is common in penny auction sites [23]. Using Figure 5, we thus filter out users who are active for less than 128 days, which we call *short-term bidders*. This leaves us with 9,105 users.

We quantify user behavior using the following features with the aim of capturing behaviors relevant to bidding performance while removing correlation between features.

- **Number of auctions.** Activity level of a user defined by number of auctions user has participated in.
- **Bids per auction.** Average number of bids placed by the user in each auction.
- **Bid response time.** Average time gap between a user's bid and the prior bid.

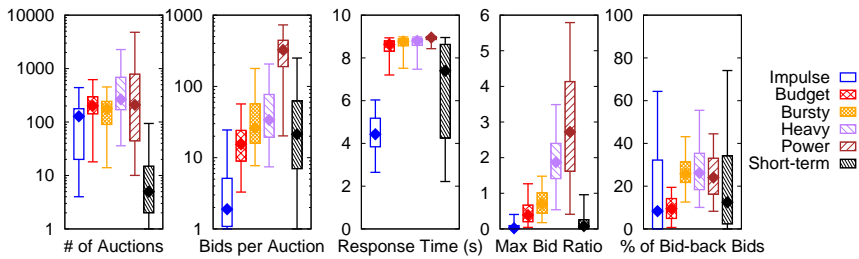


Figure 6: Distribution of feature value for each cluster and for short-term bidders. We depict each distribution with box plot quantiles (5%, 25%, 50%, 75%, 95%).

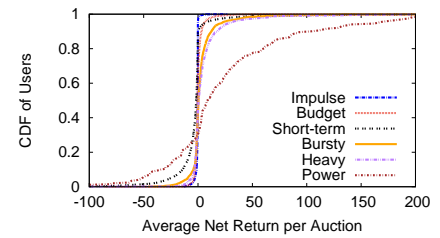


Figure 7: Distribution of earnings per auction for different types of bidders.

- **Max bid ratio.** Maximum ratio of value of bids placed against the value of the product, across all auctions participated in by the user. A value above 1 means the bidder paid more in bids than the value of the product in at least one auction.
- **Bid-back rate.** Bid-back is when a user is out-bid and she places the immediate next bid. This measures the user’s aggressiveness in auctions.

We compute the feature vector for each user (z-score normalized), and measure the Euclidean distance between user pairs. We then applied Divisive Hierarchical Clustering [14] to cluster the 9,105 active users. When breaking the users into five clusters, we achieve the best clustering quality as indicated by modularity [16].

5.2 Cluster Analysis

We manually label each cluster based on the feature values they exhibit, as shown in Figure 6. To measure performance, we show, in Figure 7, on average how much each user is winning from each auction. Each user’s net return is calculated as the total worth of items won in auctions minus the final auction price and minus the cost of bids.

- **Impulse Bidders (9.79% of users).** Contrary to common practice, these users do not wait until the last second to place their bids. Instead, their bids are placed at random points during the countdown, with a median wait time of 4.4 seconds. Most in this group failed to make money from auctions, losing a median of \$0.20 per auction.
- **Budget Bidders (55.8%).** They only place a few bids in each auction, and rarely bid back when outbid. These users are not aggressive enough, and most win less than 1% of the auctions they join. This is the user cluster with the largest median net loss, losing more than \$0.47 per auction.
- **Bursty Bidders (19.6%).** These users avoid spending too many bids in any auction, but are still able to maintain a high bid-back rate. They tend to concentrate their bids to a short period, either ending in winning the auction or running out of bids. Most users in this group produce net gains, with a median return of \$0.48 per auction.
- **Heavy Bidders (12.5%).** These users join a lot of auctions and are willing to place substantial number of bids to win the auction. Median return per auction is \$0.63.
- **Power Bidders (2.22%).** They are similar to heavy bidders, but place an order of magnitude more bids, characterized by a high max bid ratio. Given more bids placed per auction, it

is unsurprising that they are much more likely to win their auctions, with a median of 14.8% chance of winning. Median return per auction for these users is \$9.38.

By observing the clusters, we find that most DealDash bidders are actually losing money, while a small number of winners earn significant profit. The most successful users (power bidders and heavy bidders) win 13.08% and 27.00% of all auctions, despite making up less than 15% of the long-term user population. While other groups win their share of auctions, they lose far more auctions, and are generally unable to recoup their losses from the cost of bids.

6 DISCUSSION AND CONCLUSIONS

Our analysis of penny auctions and their users focused on two things, the predictability of sequential bids in auctions, and identifying common bidding behaviors in penny auctions and the results of these behaviors.

There are two takeaways from our results. First, we find that penny auctions are surprisingly predictable. Due to both the awareness of last second bidding strategies and scripts that follow them, bidding is highly periodic. In addition, users tend to exhibit highly periodic bidding behavior, which is easily captured by n-gram models, and slightly more accurately using LSTM models. Together, they suggest that the sequence of bids and timing can both be predicted with reasonable accuracy. Second, we find that most bidders tend to fall into one of a handful of clear behavioral categories, based on how patient they are, how aggressive they are, and how much money they have and are willing to use to win. We show that the large majority of users lose money, and the winnings go disproportionately to a small portion of the users (mostly power bidders willing and able to use large bid volumes to win auctions).

Finally, our results suggest that penny auctions themselves could possibly be gamed adversarially. It seems intuitive that drawing from our results, a reasonably complex model could emulate power bidder behavior, and use predictions to gain a significant advantage over its competitors and profit. We leave the development of such a model and adversarial algorithm to future work.

ACKNOWLEDGMENTS

We wish to thank our anonymous reviewers for their constructive feedback. This project was supported by NSF grants CNS-1527939 and CNS-1705042. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of any funding agencies.

REFERENCES

- [1] 2017. DealDash. (October 2017). <https://dealdash.com/>.
- [2] Corey M Angst, Ritu Agarwal, and Jason Kuruzovich. 2008. Bid or buy? Individual shopping traits as predictors of strategic exit in on-line auctions. *International Journal of Electronic Commerce* 13, 1 (2008), 59–84.
- [3] Ned Augenblick. 2015. The sunk-cost fallacy in penny auctions. *The Review of Economic Studies* 83, 1 (2015), 58–86.
- [4] Matt Backus, Thomas Blake, Dimitriy V Masterov, and Steven Tadelis. 2015. Is Sniping A Problem For Online Auction Markets?. In *Proc. of WWW*.
- [5] Ravi Bapna, Paulo Goes, Alok Gupta, and Yiwei Jin. 2004. User heterogeneity and its impact on electronic auction market design: An empirical exploration. *Mis Quarterly* (2004), 21–43.
- [6] Eric T Bradlow and Young-Hoon Park. 2007. Bayesian estimation of bid sequences in internet auctions using a generalized record-breaking model. *Marketing Science* 26, 2 (2007), 218–229.
- [7] Peter F Brown, Peter V Desouza, Robert L Mercer, Vincent J Della Pietra, and Jenifer C Lai. 1992. Class-based n-gram models of natural language. *Computational linguistics* 18, 4 (1992), 467–479.
- [8] John W Byers, Michael Mitzenmacher, and Georgios Zervas. 2010. Information asymmetries in pay-per-bid auctions. In *Proc. of ACM EC*.
- [9] Patrick Collinson. 2017. Six auction sites’ ads banned over misleading savings claims. *The Guardian*. (February 2017). <https://www.theguardian.com/money/2017/feb/22/auction-sites-ads-banned-claims-madbid-asa>.
- [10] Robert F Easley and Rafael Tenorio. 2004. Jump bidding strategies in internet auctions. *Management Science* 50, 10 (2004), 1407–1419.
- [11] Rayid Ghani. 2005. Price prediction and insurance for online auctions. In *Proc. of KDD*.
- [12] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [13] Wolfgang Jank, Galit Shmueli, and Shanshan Wang. 2006. Dynamic, real-time forecasting of online auctions via functional models. In *Proc. of KDD*.
- [14] Leonard Kaufman and Peter J Rousseeuw. 2009. *Finding groups in data: an introduction to cluster analysis*. Vol. 344. John Wiley & Sons.
- [15] David R Konkel. 2012. Costing a Pretty Penny: Online Penny Auctions Revive the Pestilence of Unregulated Lotteries. *Seattle University Law Review* 36 (2012), 1967.
- [16] Mark EJ Newman. 2006. Modularity and community structure in networks. *PNAS* 103, 23 (2006), 8577–8582.
- [17] Young-Hoon Park and Eric T Bradlow. 2005. An integrated model for bidding behavior in Internet auctions: Whether, who, when, and how much. *Journal of Marketing Research* 42, 4 (2005), 470–482.
- [18] Brennan C Platt, Joseph Price, and Henry Tappen. 2013. The role of risk preferences in pay-to-bid auctions. *Management Science* 59, 9 (2013), 2117–2134.
- [19] Alvin E Roth and Axel Ockenfels. 2002. Last-minute bidding and the rules for ending second-price auctions: Evidence from eBay and Amazon auctions on the Internet. *American Economic Review* 92, 4 (2002), 1093–1103.
- [20] Jarrod Trevathan and Wayne Read. 2009. Detecting shill bidding in online English auctions. *Handbook of research on social and organizational liabilities in information security* 46 (2009), 446–470.
- [21] Dennis Van Heijst, Rob Potharst, and Michiel van Wezel. 2008. A support system for predicting eBay end prices. *Decision Support Systems* 44, 4 (2008), 970–982.
- [22] Shanshan Wang, Wolfgang Jank, and Galit Shmueli. 2008. Explaining and forecasting online auction prices and their dynamics using functional data analysis. *Journal of Business & Economic Statistics* 26, 2 (2008), 144–160.
- [23] Zhongmin Wang and Minbo Xu. 2016. Selling a dollar for more than a dollar? Evidence from online penny auctions. *Information Economics and Policy* 36 (2016), 53–68.