QUANTIFYING THE ECONOMIC BENEFITS OF EFFECTIVE REDRESS: LARGE E-COMMERCE DATA SETS AND THE COST-BENEFIT CASE FOR INVESTING IN DISPUTE RESOLUTION

Colin Rule[*]

SUMMARY

One of the biggest challenges for dispute resolution has been demonstrating the economic benefit of investments in fair and effective redress systems. Many studies have demonstrated that Alternative Dispute Resolution (ADR) delivers high satisfaction scores, but user-reported satisfaction can be an unreliable metric to measure success. Plus it is difficult to translate satisfaction scores into specific economic benefit. In this article, the author uses results culled from very large e-commerce data sets to demonstrate the quantifiable benefit of investments in effective dispute resolution processes. This data is not based on user-reported satisfaction, but on an analysis of the actual behavior of users before and after a dispute event. These results clearly show the value of investments in dispute resolution and offer hard evidence of economic benefits that can be gleaned from the deployment of effective redress processes. These results are relevant not only to e-commerce service providers, but to any organization that interacts directly with customers, from private companies to public agencies.

I. INTRODUCTION

Over the past several decades the Alternative Dispute Resolution (ADR) field has struggled to quantitatively demonstrate the benefit of the work that it does.[1] Many studies have shown dispute resolution services often receive very high satisfaction scores from users.[2] However, the difficulty has been in demonstrating that this improved satisfaction generates

* CEO of Modria.com, an ODR provider based in Silicon Valley, and author of *Online Dispute Resolution for Business* (Jossey-Bass, 2002). From 2003 to 2011, he was Director of Online Dispute Resolution for eBay and PayPal. A long time dispute resolution mediator, trainer, and consultant, he is currently Co-Chair of the Advisory Board of the National Center for Technology and Dispute Resolution at UMass-Amherst and a Non-Resident Fellow at the Center for Internet and Society at Stanford Law School.

1. *See* Frank E.A. Sander, *The Future of ADR the Earl F. Nelson Memorial Lecture*, 2000 J. DISP. RESOL. 3, 5–6 (2000).

2. Frank E.A. Sander & Lukasz Rozdeiczer, *Matching Cases and Dispute Resolution Procedures: Detailed Analysis Leading to a Mediation-Centered Approach*, 11 HARV. NEGOT. L. REV. 1, 14 (2006).

concrete and replicable economic benefit.[3] This has made the challenge of arguing for continued investment in dispute resolution systems quite difficult. Several high quality dispute resolution programs have fallen victim to budget cuts[4] because they were unable to justify continued expenditure on their operations from a purely cost-benefit perspective.

From 2003 to 2011, I was fortunate enough to serve as the first Director of Online Dispute Resolution at one of the largest e-commerce marketplaces in the world. During that time I was asked to build and implement online dispute resolution services from scratch for a global community of more than 200 million users. The resolution systems my team designed and launched now handle more than 60 million disputes per year, a case volume many times larger than the U.S. court system.[5]

One of the challenges I encountered when I arrived at the company was explaining the benefit of investment in online dispute resolution system to executives and decision makers who did not see dispute resolution as a self-evident good. Many of the individuals responsible for guiding the strategic direction of the company were focused primarily on profit and loss, not user satisfaction. Every new project had to struggle to secure its share of scarce development resources. As part of each product proposal the author had to conduct a cost-benefit analysis for the project and indicate the net-present-value of following through with the initiative. If the cost to complete the project exceeded the projected financial benefits (either from increased revenues or cost savings), there was little chance the project would be approved. As a result, I found myself for the first time in a situation where I had to explain the benefits of dispute resolution processes purely in terms of dollars and cents. I knew that making this argument had been a struggle for other dispute resolution initiatives, and that satisfaction-based justifications had a mixed record of success, so I decided to take a new approach.

Being part of a well-resourced technology company and working with a very large user base enabled me to have some unique analytic opportunities. To monitor the continuing operation of resolutions on the site I was given access to a giant database called the "data warehouse" which was

---

3. *See* Sander, *supra* note 2, at 6.

4. Denise Richardson, *Local Dispute Center Loses Funds, Jobs*, THEDAILYSTAR, http://thedailystar.com/localnews/x1678756287/Local-dispute-center-loses-funds/jobs/print (last visited May 16, 2012).

5. In a twelve-month period from March 2010 to March 2011, 294,336 civil cases were filed in U.S. district courts. STATISTICS DIVISION, ADMIN. OFFICE OF THE U.S. COURTS, FEDERAL JUDICIAL CASELOAD STATISTICS: MARCH 31, 2011, Table C (2011) *available at* http://www.uscourts.gov/Viewer.aspx?doc=/uscourts/Statistics/FederalJudicialCaseloadStatistics/2011/tables/C00Mar11.pdf.

filled with petabytes[6] of information tracking the behavior of all of the site's users over several years. The information in the data warehouse was extremely specific; it tracked every click, purchase, and visit of every user over a multiyear period. Nothing was forgotten, and it was all available to be queried in real time. It seemed to me that this data could help me make the case for investing in dispute resolution.

II. THE LIMITATIONS OF SATISFACTION AS A METRIC FOR SUCCESS

At my company, user satisfaction was and still is a top priority. As a result, we had very sophisticated systems to track both user satisfaction regarding individual processes and flows, as well as the user's satisfaction regarding their overall relationship with the site. I actually led my division's effort to interpret these results, and we employed a unified industry-standard metric—called NPS, or Net Promoter Score[7]—across all of our surveys to track our progress in improving satisfaction over time. What we discovered in closely analyzing that data over several years was that satisfaction as measured through self-reported surveys is a very imprecise way to measure success.

Our analysis revealed that satisfaction was often correlated directly with outcome, so users that got what they wanted (or "won") in a dispute were satisfied, and users who "lost" were dissatisfied. In addition, our results were warped by flows where we wanted users to be dissatisfied. For instance, if we were limiting or closing a user's account due to intentional and repeated policy violations, we had no problem when the user indicated they did not like the experience. We also found it was relatively easy to manipulate satisfaction numbers, simply by automatically paying out more claims or "no faulting" cases so no one lost. Those short term changes might improve satisfaction for a limited period of time, but they were not based on concrete changes in our platform or service. As a result, the benefits would disappear quickly if the payments stopped.

Another shortcoming involved the self-reported nature of satisfaction. Users would sometimes insist that a certain new feature would massively improve their satisfaction, but when we launched it, it had little effect. Or conversely, we would suggest a new feature that users indicated they were indifferent about, but when it launched it would generate a major increase in satisfaction. Finally, users often misreported how satisfied they were or how that satisfaction affected their usage of our site. They might say they were

---

6. A petabyte is a unit of measurement approximately equal to one million gigabytes or 1000 terabytes. *See* Philip Beatty, *The Genesis of the Information Technologist-Attorney in the Era of Electronic Discovery*, 13 J. TECH. L. & POL'Y 261, 288 n.87 (2008).

7. For information on Net Promoter Score, see NET PROMOTER, http://www.netpromoter.com/np/calculate.jsp (last visited May 1, 2012).

going to close their account and never come back as the result of a bad experience, but two days later they would be back on the site, using our services more than they had ever used them before. We found it was very difficult to use satisfaction metrics to guide our strategic decision making because they were often so disconnected from reality.

III. RELYING ON USAGE DATA INSTEAD OF SATISFACTION DATA

What we found to be a better indicator of the real impact of our decisions was the detailed data we captured in the data warehouse. Instead of relying on surveys to determine self-reported satisfaction, we could query the database to determine exactly how each user's behavior changed after they had a particular experience on the site. In many respects, the data warehouse had a better understanding of the user's satisfaction than the users themselves had. A customer might say on a survey that a particular experience soured them on the site and they decreased their usage afterward, but the information in the data warehouse told the true story.

My team decided to conduct our own investigation into the usage metrics of the individual users going through our online dispute resolution systems. We suspected that the information available in the data warehouse would tell a different story about the specific impacts of our resolution processes. We also believed that the data warehouse would enable us to examine a larger and more detailed data set than any study feasible in the off-line environment, which might lead to some interesting insights that could be extrapolated beyond our specific flows.

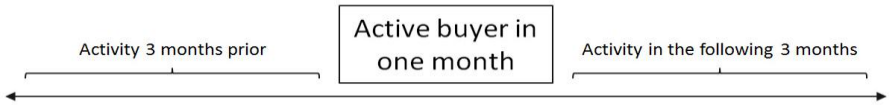
IV. DESIGNING THE EXPERIMENT

The very smart people in the data analytics team came up with an approach that they thought could best get at the nub of the question. By examining a very large sample of users in our data warehouse, many hundreds of thousands of individual accounts, we could structure a backward-looking A/B test[8] between two separate pools of accounts to tease out the specific impact of going through the online dispute resolution flows we had built on the site.

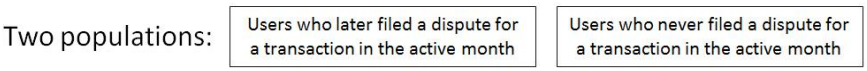
---

8. An A/B Test is a controlled experiment where "users are randomly exposed to one of two variants: Control (A), or Treatment (B) . . . . If the experiment was designed and executed properly, the only thing consistently different between the two variants is the change between the Control and Treatment, so any differences in the [Overall Evaluation Criterion] are inevitably the result of this assignment, establishing causality." Ron Kohavi, et al., *Controlled Experiments on the Web: Survey and Practical Guide*, 18 DATA MINING & KNOWLEDGE DISCOVERY 140, 149 (2009) *available at* http://www.springerlink.com/content/r28m75k77u145115/fulltext.pdf.

We decided to focus on active buyer accounts in a particular month. We would then analyze the activity of each account three months prior to the month in question and three months after the month in question:



We would then split that set of accounts into two separate populations: one pool of users who filed a dispute in the active month in question, and another pool of users who did not file a dispute in that month:



Next, we would then generate an Activity Ratio for each account, indicating how active each buyer was on the site for the test periods. This ratio would be calculated by dividing the buyer's Total Payments Volume for the three months post by the buyer's Total Payments Volumes in the three months prior:

$$\text{Activity Ratio} = \frac{\text{Total Payments Volume 3 months after}}{\text{Total Payments Volume 3 months prior}}$$

The beauty of this metric was that it represented the actual impact of going through the online dispute resolution process on the user's behavior, not the buyer's perception of how going through the process impacted their behavior. The other positive aspect of this metric was that it was easily convertible into actual economic benefit to the company because Total Payment Volume can be converted into estimated profits. We knew the exact percentage of profit for every additional dollar of Total Payment Volume processed, so this made the cost-benefit calculation much simpler.

Our hypothesis was that we would find similar results to the dynamics we had witnessed on the satisfaction surveys: activity on the site after a dispute event would be correlated to dispute outcome. Most of us believed that increases in activity were likely to directly correlate to positive outcomes in the dispute resolution process. For example, if a buyer reported a dispute and ended up receiving a full refund, the presumption was that because they achieved a positive outcome they would increase their activity on the site moving forward. The flip side to that assumption was that if a buyer was to receive a less than favorable outcome, such as the decision that the seller was in the right and that the buyer was not entitled to a refund, they would

be angry and there would be a corresponding drop in their activity moving forward. So "winning" buyers (e.g., buyers who received a full refund) would increase their usage, and "losing" buyers (e.g., buyers who did not get a refund) would decrease their use of the site. We all suspected that was most likely what the data would demonstrate.

V. WHAT WE DISCOVERED

To our surprise, once we collected the data, we found that it did not validate this hypothesis. The results revealed a clear benefit to user activity as a result of engaging in the online dispute resolution process. However, the results also demonstrated that, on average, users who reported a transaction problem and went through the online dispute resolution process increased their usage of the marketplace, regardless of outcome.

What that meant was that buyers who "won" their case increased their activity, but buyers who "lost" their case also increased their activity. Now it is true that the buyers who lost their case did increase their activity at a slower rate than the buyers who won their case, but most surprisingly, both of those buyers increased their activity more than buyers who never filed a dispute in the first place.

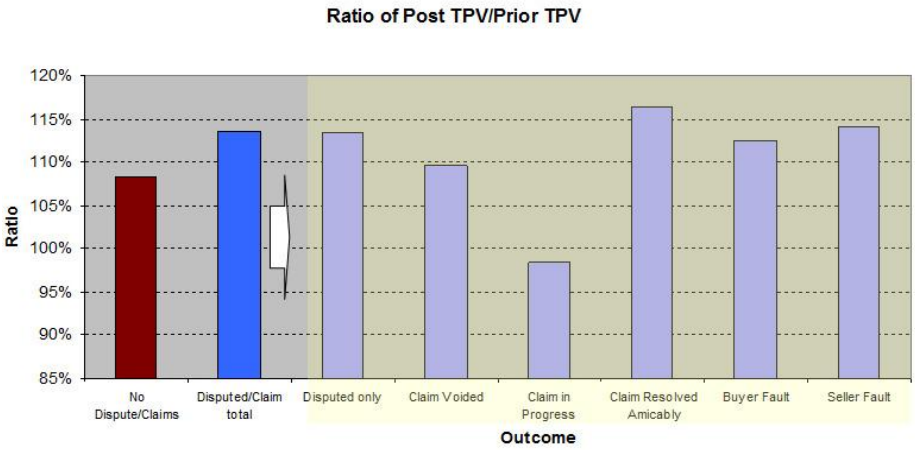


Figure A: Activity Ratios of Buyer Accounts by Outcome Achieved

As indicated in Figure A, the Activity Ratio for buyers who did not file a dispute in the active month was about 108%, and the Activity Ratio for buyers who did file a dispute in the active month was about 114%. But what was more interesting was the fact that every outcome of the dispute process had a higher Activity Ratio than the non-filing buyers, even when the claim was voided or the buyer was found to be at fault. The group of buyers who had the highest post-dispute Activity Ratio was the group of those buyers

who had their claims resolved amicably, through mutual agreement with their sellers. This group had an Activity Ratio of approximately 117%, higher even than the buyers who won their claims outright (114%).

The only group of buyers who filed a dispute and decreased their activity on the site in the three months after the active month were buyers for whom the resolution process took a very long time (identified as "Claim in Progress" in Figure A). These buyers filed a dispute and, for one reason or another, had the resolution of that dispute take longer than six weeks. If the dispute was resolved within six weeks, then the Activity Ratio was higher than the non-dispute-filing accounts, but if the resolution process stretched beyond six weeks, then the Activity Ratio fell lower than the non-filing accounts. That was the only outcome in which the Activity Ratio was lower than the non-filing buyers. However, as you can see in Figure B, that group of buyer accounts was less than 1% of the overall pool of accounts observed:

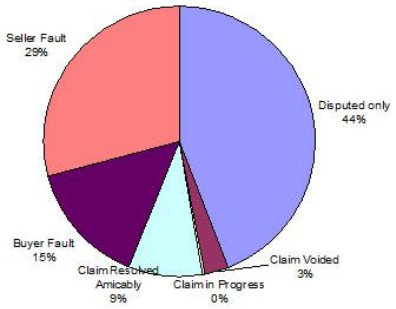


Figure B: Account Distribution of Outcome

Another interesting result we discovered was that these benefits held for filing buyer accounts across all activity levels, and that these benefits are statistically significant. In Figure C you can see that the accounts that filed a dispute had a higher Activity Ratio regardless of whether the buyer spent $100 per month or more than $1000 per month. The error bars indicate the upper and lower bound of the ratio at the 95% confidence level.

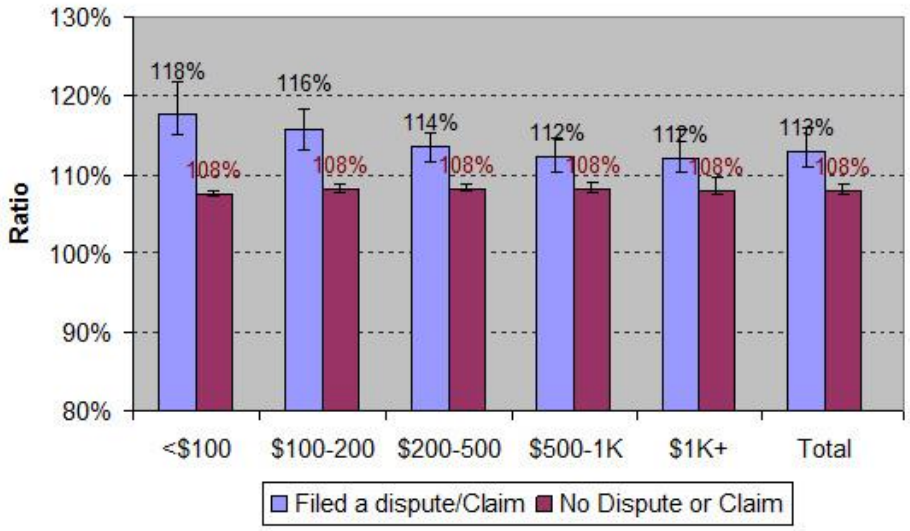


Figure C: Account Distribution of Outcome

We also found that the activity benefit was much more pronounced for buyer accounts where the dispute filed during the active month was the first dispute ever filed for that account. For buyer accounts with previous dispute experience the impact was far more muted. Additionally, we found that the positive impact on activity went away if the buyer turned instead to their credit card issuer for redress as opposed to the on-site resolution flow we provided.

VI. EXPLAINING THE RESULTS

This data was quite revolutionary to many, as it contradicted several long held beliefs within the company. Some executives were incredulous when they first reviewed the results, sarcastically suggesting that maybe we should intentionally give all our new buyers a dispute in their first year on the site so as to push up their activity.

But upon reflection, the results made a lot of sense. It is well understood that user trust is a crucial driver to growing usage of online services.[9] Resolution is a core component of user trust.[10] Many new users may have

9. *See, e.g.*, Lucille M. Ponte, *Boosting Consumer Confidence in E-Business: Recommendations for Establishing Fair and effective Dispute Resolution Programs for B2C Online Transactions*, 12 ALB. L.J. SCI. & TECH. 441, 443 n.3 (2002).

10. "Dispute resolution processes are generally perceived as having a single function, that of settling problems. What has come to be understood online, perhaps more than it is offline, is that dispute resolution processes have a dual role, that of settling disputes and also

doubt that if they encounter a transaction problem they will be able to get it resolved quickly and effectively. This lack of confidence acts as a break on their usage of the service in question. This data made clear that that once a buyer gained a first-hand understanding of the available resolution options, and that effective systems were available to help them resolve any transaction problems they encountered, that understanding would encourage them to increase their usage of the service over time.

As an example, imagine you were buying gifts for your family and friends over several weeks leading up to the holiday season. As gift ideas come to you, you may end up purchasing items across a variety of different online marketplaces. Initially you may have little preference as to which marketplace you use for each item. As the packages arrive in the mail, you may even forget which items you purchased from which marketplaces. And as the recipients open their gifts, all you know is that you bought the items online and they arrived without a problem. Buyers often have high expectations for their online purchases,[11] so when everything goes smoothly the transaction (and the transaction environment) often makes little impression on the purchaser.

But imagine that one of the items arrives and there is a problem. Maybe it was damaged in shipping, or maybe the wrong item was delivered. When that happens, you as the purchaser must pay individual attention to that particular transaction. You go back to your e-mail, search for the item receipt, and determine which marketplace the item was purchased from. Then you go to the website of the marketplace and try to determine what you need to do to get the problem resolved. That is the moment at which the buyer experiences an in-depth and unexpected interaction with an e-commerce marketplace. That is the moment where loyalty is imprinted. If the marketplace provides an easy to find process for resolving the problem, a strong impression is made in the mind of the buyer. If the marketplace does not provide any easy to discover process for resolving the problem, the buyer's experience instead is one of frustration, which creates a strong impression in the other direction.

The results of this research demonstrate that once a buyer goes through the online dispute resolution process, learning both how to initiate it and how it delivers fair and efficient resolutions, that awareness drives them to increase their use of the overall website by a statistically significant amount.

---

of building trust." Ethan Katsh, *Online Dispute Resolution: Some Implications for the Emergence of Law in Cyberspace*, 10 LEX ELECTRONICA, no.2, Winter 2006, at 1, 6 (available at http://www.lex-electronica.org/articles/v10-3/katsh.htm) (last visited May 16, 2012).

11. Kristina Knight, *Online Shopping Expectations are Rising*, BIZREPORT.COM, http://www.bizreport.com/2008/01/online_shopping_expectations_are_rising.html# (last visited May 16, 2012).

The group of buyers who increased their activity on the site the most in the wake of a dispute filing was the group of buyers who achieved an amicable resolution to their dispute. The Activity Ratio of these buyers showed the greatest increase, even greater than those users who "won" their case and received a full refund as mandated by the marketplace administrator. We believe the explanation for this phenomenon is that trust in your fellow users to do the right thing in good faith is more powerful than the belief that a marketplace administrator will intervene and use their power to decide disputes between users who disagree. Having a transaction partner hear your complaint and resolve the issue is a much more effective trust building outcome than relying upon a site administrator to mete out justice in each case.

The only buyers who decreased their activity after filing their first dispute were buyers for whom the process took a long time, more than six weeks. This lesson affirmed feedback we had heard previously indicating that buyers preferred to lose their case quickly rather than have the resolution process go on for an extended period of time. The frustration associated with a long resolution process outweighed the benefit from getting a positive outcome because buyers value their time more than the money in question when it comes to low dollar value transactions. The buyers in these cases learn a different lesson: the marketplace does not have a quick and effective resolution process in place, and that realization displaces any benefit that comes from educating buyers about the existence of online dispute resolution.

VII. RAMIFICATIONS OF THIS STUDY

The results of this study demonstrate the clear and quantifiable economic benefit to organizations that invest in the development and promotion of online dispute resolution tools and systems. Many businesses have traditionally downplayed their resolution systems because they thought that talking about problems with buyers would make them less likely to utilize the services of the website in question. Often resolution processes are hidden deep inside help content, or made available only to users who proactively contact the website to complain about a problem. The results of this research demonstrate the short sightedness of that approach.

Problem resolution is perhaps the most important loyalty moment for consumers. In e-commerce, providing transactions with no problems, where the buyer purchases an item and it arrives without a hitch, is the default expectation—the table stakes for e-commerce, if you will. The loyalty moment comes when the buyer experiences something out of the ordinary. That is when the marketplace is presented with an opportunity to step up, do the right thing, and make a lasting, positive impression on that customer.

Some e-commerce companies have understood this dynamic intuitively for many years. The experience of those companies has demonstrated results

consistent with the conclusions from our study. Those websites often provide very streamlined resolution processes to their buyers, along with very generous policies around refunds and returns. Over time those services have claimed a greater share of the overall e-commerce marketplace.[12] Other services that have demonstrated less commitment to streamlined problem resolution may deliver better results over a shorter period of time, but over the long term their share of e-commerce activity has fallen in comparison to the firms that have understood and embraced the importance of quick and effective problem resolution.

This is true not only for e-commerce companies, but for all organizations that provide client services, from cell phone companies to airlines to government agencies. Being up front with customers about resolution systems, and providing an excellent resolution experience once a problem crops up, is essential to building loyalty with that customer. What this data demonstrates clearly is that those investments generate real returns which more than compensate for the expense of putting those systems in place. Proactive communications to customers and clients about resolutions is good business, and organizations that ignore or downplay resolutions do so at their own peril.

12. *See* Henry Blodget, *Amazon is Still Eating eBay's Lunch*, BUSINESSINSIDER, http://articles.businessinsider.com/2010-01-15/tech/30086371_1_powersellers-ebay-share (last visited May 16, 2012); Rob Enderle, *EBay vs. Amazon: An Interesting Lesson in Customer Care*, ITBUSINESSEDGE, http://www.itbusinessedge.com/cm/blogs/enderle/ebay-vs-amazon-an-interesting-lesson-in-customer-care/?cs=49557&page=2 (last visited May 16, 2012); *Amazon.com Price vs. eBay Price*, AMAZONWEBSERVICES.COM, http://node_charts_production.s3.amazonaws.com/56fd95dfec274958c3f229941cf676a7.png (last visited May 16, 2012).