American Business Law Journal
Volume 58, Issue 2, 225–269, Summer 2021

Does Conjoint Analysis Reliably Value Patents?

Bernard Chao and Sydney Donovan***

Modern technology products are often covered by thousands of patents. Yet awards for a single component have averaged a surprisingly high 9.98 percent of the infringing product's price. To curb such disproportionate awards, the law insists that damages reflect the contribution made by the patent. But determining how to apportion damages in this way has proved to be elusive. One emerging technique that appears to offer rigor is conjoint analysis, a type of survey borrowed from the marketing world. This article explores the validity of the conjoint analysis technique by running two conjoint analysis surveys. Unfortunately, we found serious problems. First, the results of our surveys yielded irrationally high numbers. Most survey features suffered from bizarrely high valuations. Second, we demonstrate how experts can manipulate the results by selecting among a number of different ostensibly reasonable statistical choices and picking the one that yields the most desirable outcome. Based on these findings, we provide several recommendations. First, we argue that courts should not allow evidence of conjoint analysis to show the monetary value of specific features. However, we recognize that there is support for using conjoint analysis to provide relative valuations (i.e., feature A is worth significantly more than feature B). To the extent that courts permit this use, we suggest ways to ensure that experts employ the best science available. These recommendations include assuring that experts accurately depict variability in their results and requiring experts to "preregister" the approach they intend to use with the court.

*Bernard Chao is a Professor at the University of Denver Sturm College of Law. Thanks to David Abrams, Hillel Bavli, Suneal Bedi, John Bliss, Tom Cotter, Michael Frakes, Doug Kidder, Mark Lemley, Amy Landers, Brian Love, Christopher Robertson, Samantha Zyontz, and participants of the 2020 Works in Progress IP Colloquium and 2020 Intellectual Property Scholars Conference for all their helpful comments. Thanks to Catherine Durso for providing advice on our statistical analyses and Michelle Penn for research assistance. Funding for the surveys was provided by the Hughes-Ruud endowment of the University of Denver Sturm College of Law.
**Sydney Donovan is an associate with WilmerHale. This article was written while Ms. Donovan was a student at the University of Denver, and the contents do not represent the views of WilmerHale or its clients.

Introduction

Modern complex products are covered by tens or even hundreds of thousands of different patents.[1] For example, smartphones contain countless patented features that include the graphical user interface, microprocessor, memory chips, and communication protocols.[2] Patent lawsuits typically involve only one patent, or in more complex cases, a handful of patents.[3] Given these numbers, one might expect damages in patent cases to be a small fraction of product price. This expectation turns out to be far from reality. A 2007 study found that reasonable royalty awards for a single component in a complex multicomponent product averaged 9.98 percent of the infringing product's price.[4] In response to such disproportionately high damage awards, the Federal Circuit has repeatedly insisted that courts somehow apportion damages so that awards reflect only the contribution made by the patent and exclude value from other inputs.[5]

In an effort to satisfy the apportionment requirement, attorneys have offered a variety of crude estimates to calculate the value a specific patent contributes. Representative approaches include looking at the proportion of area the patented invention occupies on a semiconductor chip, counting lines of code, counting features, and even looking at the number of pages found in a published standard.[6] One emerging technique

---

[1] Patent Remedies and Complex Products: Toward a Global Consensus 1 (C. Bradford Biddle, Jorge L. Contreras, Brian J. Love & Norman V. Siebrasse eds., Cambridge U. Press 2019).

[2] *See generally* Ann Armstrong, Joseph J. Mueller & Timothy D. Syrett, *The Smartphone Royalty Stack: Surveying Royalty Demands for the Components Within Modern Smartphones* (working paper), https://www.wilmerhale.com/-/media/files/shared_content/editorial/publications/documents/the-smartphone-royalty-stack-armstrong-mueller-syrett.pdf.

[3] *See generally* Fish & Richardson, *A Guide to Patent Litigation in Federal Court*, https://www.fr.com/wp-content/uploads/2019/05/2019-Q2-Guide-to-Patent-Litigation-in-Federal-Court-final.pdf.

[4] Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 Tex. L. Rev. 1991, 2032 tbl.1 (2007).

[5] *See, e.g.*, Virnetx, Inc. v. Cisco Sys., Inc., 767 F.3d 1308, 1326 (Fed. Cir. 2014).

[6] *See* Damien Geradin & Ann Layne-Farrar, *Patent Value Apportionment Rules for Complex, Multi-Patent Products*, 27 Santa Clara High Tech. L.J. 763, 777–86 (2011) (discussing various methods for apportioning the value of individual patents to multi-patent products).

that appears to promise more rigor is conjoint analysis, a type of survey commonly used in the marketing world.[7] However, there has been very little scrutiny of conjoint analysis in the legal context.[8]

This article explores the validity of the technique by running two conjoint analysis surveys and finds serious problems with the technique. Our chief finding is that conjoint analysis yields inaccurate valuations, mostly producing irrationally high dollar values. These findings are consistent with what others have previously reported.[9] Accordingly, we recommend that courts not permit the use of conjoint analysis to provide specific monetary valuations.[10] However, we recognize that conjoint analysis may still be useful for determining relative valuations (i.e., to suggest that a given feature is worth more or less than other features). Even in this context, we urge caution. Surveys are useful only to the extent that respondents understand the choices that they make, which can be problematic when the patented features are complex. Moreover, as with many scientific techniques, the results of a conjoint analysis survey can be manipulated. We demonstrate how experts can manipulate their results to favor one side. Finally, to the extent that courts permit the use of conjoint analysis, we provide several recommendations to curb these kinds of abuse.

Part I provides a basic background of the law of patent damages and discusses different approaches experts have been using to isolate the value of a patented feature in modern multicomponent products. Conjoint analysis is a relatively new approach that, at least superficially, looks to be more rigorous than prior approximations. Part II describes the conjoint analysis survey and how experts use the resulting data to calculate average willingness to pay (WTP) values for specific features. Part II then describes some current critiques of this approach.

Part III describes both a preliminary survey on smartphone features and a pair of conjoint analysis surveys. The preliminary survey simply

---

[7]*See What is Conjoint Analysis?*, CONJOINT.LY, https://conjointly.com/guides/what-is-conjoint-analysis/ (last visited Apr. 8, 2021).

[8]*See* Jonah M. Knobler & Joshua Kipnees, *Conjoint Analysis: No Silver Bullet for Calculating Class-Wide Damages*, PATTERSON BELKNAP: MISBRANDED (Mar. 1, 2019), https://www.pbwt.com/misbranded/conjoint-analysis-no-silver-bullet-for-calculating-class-wide-damages).

[9]*See infra* notes 37–41 and accompanying text.

[10]*See* Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 593–94 (1993) (requiring that expert testimony concern scientific knowledge and explaining that one consideration of that threshold is the "potential rate of error.").

asked respondents to rank the value of six smartphone features. The results of the preliminary survey informed the design of the two conjoint analysis surveys, one with less valuable features and one with more valuable features. In the first conjoint analysis survey, we included two less valuable features (noise reduction and water resistance) with the two primary features of interest (battery life and location privacy). In the second conjoint analysis survey, we included two more valuable features (storage and screen resolution) with the primary features of interest.

Part IV reports on the results of the two conjoint analysis surveys. First, we show the two surveys yielded disproportionately high WTP values. Second, we demonstrate how using average price sensitivity to calculate each respondent's WTP, instead of individual price sensitivity, can artificially reduce variability and make the results appear more consistent than they truly are. Third, we demonstrate how using different trimming techniques affects our results, and how such trimming techniques make it possible for experts to select an approach that favors their side, a practice sometimes called "data fishing."[11] Finally, we discuss our hypothesis that feature values could be manipulated by including less valuable "other features" in the survey design. While some results suggest the hypothesis might be true, others do not, and our results should be considered inconclusive.

Part V provides several recommendations with accompanying justifications. First, we recommend that courts not permit the use of conjoint analysis to provide specific monetary valuations but allow its use only to provide relative valuations in certain contexts. Second, to the extent that courts allow the use of conjoint analysis, we recommend that they require experts to use individual price sensitivities (and not average price sensitivities) when calculating individual willingness to pay values. Third, we recommend that courts require experts to testify about the confidence intervals associated with their findings in order to give fact finders an understanding of the full range of likely outcomes. Finally, to avoid manipulation, we recommend that courts require experts to file a report explaining the expert's survey design and statistical approach prior to running the conjoint analysis survey.

[11] *See* Jepchirchir Kibet Lorna, *Analysis of Demand Based Forecasting Methods Employed in Kerio Valley Development Authority, Kenya*, 6 Global J. of Advanced Res. 236, 239 (June 30, 2019), http://www.gjar.org/publishpaper/vol6issue6/d897r87.pdf.

I. Patent Damages

Under 35 U.S.C. section 284, patent holders in the United States are entitled to recover lost profits or reasonable royalties.[12] Reasonable royalty awards are the more common form of damages awarded.[13] For patent holders that do not compete in the marketplace, their only choice is to recover reasonable royalties. However, even practicing entities recover reasonable royalties in 81 percent of cases with damage awards.[14] The law instructs fact finders to determine reasonable royalty awards based on the so-called "hypothetical negotiation."[15] In other words, reasonable royalties should represent the royalty that a willing licensor and willing licensee would have agreed to at the time of first infringement. For years, this determination was made by relying on the fifteen *Georgia Pacific* factors.[16] Many commentators have criticized this test as so complicated that it inadvertently gives juries unfettered discretion to decide damages.[17] As a result, some courts have begun to simplify the test.[18]

Regardless of how the hypothetical negotiation is framed, the Federal Circuit has emphasized that awards must consider apportionment princi-

---

[12]35 U.S.C. § 284 (2020).

[13]*See* Biddle *et al., supra* note 1, at 6.

[14]Landan Ansell, Ronen Arad, Doug Branch, HyeYun Lee, Adil Pasha & Paul Robinson, 2018 Patent Litigation Study, PwC 6 (2018), https://www.pwc.com/us/en/forensic-services/publications/assets/2018-pwc-patent-litigation-study.pdf (noting that for practicing entities, 60 percent of cases involved only reasonable royalties while 21 percent involved both lost profits and reasonable royalties).

[15]*See, e.g.*, ResQNet.com, Inc. v. Lansa, Inc., 594 F.3d 860, 868 (Fed. Cir. 2010).

[16]Georgia-Pacific Corp. v. U.S. Plywood Corp., 318 F. Supp. 1116, 1120 (S.D.N.Y. 1970), *modified sub nom*. Georgia-Pacific Corp. v. U.S. Plywood-Champion Papers, Inc., 446 F.2d 295 (2d Cir. 1971).

[17]Daralyn J. Durie & Mark A. Lemley, *A Structured Approach to Calculating Reasonable Royalties*, 14 Lewis & Clark L. Rev. 627, 632–33 (2010) (arguing that the complexity of the current fifteen factor test for determining reasonably royalties allows judges to "simply give up and defer to whatever the jury awards").

[18]For example, the Northern District of California's model patent juries were updated in October 2019 and no longer include the *Georgia-Pacific* factors. *See* U.S. Dist. Ct. N. Dist. of Cal., Model Patent Jury Instructions 49 (2019), https://www.cand.uscourts.gov/forms/jury-instructions/.

ples.[19] Apportionment reflects the idea that a reasonable royalty should somehow consider only the value that the patented invention contributes to the infringing product and exclude other contributions (e.g., other patents, the infringer's marketing efforts, and risk taking). While the concept of apportionment is both simple and sensible, it is not clear how to ensure (1) that parties and their experts use apportionment principles in calculating their damage demands and (2) that juries properly take apportionment into account in their final awards.[20]

Parties have tried numerous approaches. Some court decisions seem to assume that relying on roughly comparable licenses to establish royalty rates somehow implicitly incorporates apportionment principles.[21] In addition, attorneys have sought to calculate explicitly the contribution a patent makes to a product. In theory, any approach that relies on an expert must satisfy Federal Rule of Evidence 702, which requires that expert testimony be the product of "reliable principles and methods."[22] In *Daubert v. Merrell Dow Pharms.*, the Supreme Court interpreted this standard to mean that the methodology underlying any testimony must be "scientifically valid."[23] Although the Court did not provide a definitive test, it discussed several factors that inform this

---

[19]Virnetx, Inc. v. Cisco Sys., Inc., 767 F.3d 1308, 1328 (Fed. Cir. 2014) ("[T]he district court should have exercised its gatekeeping authority to ensure that only theories comporting with settled principles of apportionment were allowed to reach the jury."); Ericsson, Inc. v. D–Link Sys., Inc., 773 F.3d 1201, 1226 (Fed. Cir. 2014) ("When the accused infringing products have both patented and unpatented features, measuring this value requires a determination of the value added by such features."). This requirement can be traced back to *Garretson v. Clark*, 111 U.S. 120, 121 (1884) (saying that the patentee must give evidence apportionment in every case).

[20]*See generally* Bernard Chao, *Implementing Apportionment*, 2019 Patently-O Pat. L.J. 20 (2019) (discussing the difficulties of implementing apportionment).

[21]Elbit Sys. Land & C4I Ltd. v. Hughes Network Sys., LLC, 927 F.3d 1292, 1301 (Fed. Cir. 2019) (by using existing license as a "starting point," apportionment was "already built in."); Sprint Commc'ns Co. v. Time Warner Cable, Inc., 760 F. App'x 977, 983 (Fed. Cir. 2019) (finding that the use of two prior licenses and a prior damages verdict "provides strong support" to show that the award "satisfied the requirement of apportionment.").

[22]Fed. R. Evid. 702.

[23]*Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 592–93 (1993).

analysis. Judges should consider whether a given approach "can be (and has been) tested"; whether it has been subject to peer review and publication; what the potential error rate might be; and whether the method has "general acceptance."[24]

Several recent approaches to apportion patent damages do not appear to satisfy this standard. Courts have permitted experts to use several types of back-of-the-envelope calculations including the die area method, source code method, and feature count method.[25] These techniques all involve counting the amount of a particular item related to the accused infringing feature and then comparing it to the total amount of that item in the entire infringing product. For example, in the die area method, the amount of die area occupied by the accused feature is compared with the total amount of die area occupied by the infringing semiconductor chip.[26] While these techniques may all be reasonable, rough estimates, none of them could be characterized as "scientifically valid" as required by *Daubert v. Merrell Dow Pharms*.

While some courts have allowed the parties to use each of these approaches to help apportion patent damages, the viability of these approaches is uncertain. In a patent lawsuit brought by the California Institute of Technology, Apple sought to have its experts apportion damages by relying on the three approaches discussed above as well as a fourth approach: counting pages of the 802.11 Wi-Fi standard as they corresponded to the accused technology and comparing that number to the total number of pages in the standard.[27]

---

[24]*Id*. at 593–95.

[25]Ziilabs Inc., Ltd. v. Samsung Elecs. Co., No. 2:14-CV-203-JRG-RSP, 2015 WL 11110651, at *5 (E.D. Tex. Dec. 4, 2015) (permitting patentee's expert to consider the percentage surface area on a semiconductor die for an apportionment analysis); Lucent Technologies, Inc. v. Gateway, Inc., 580 F.3d 1301, 1333 (Fed. Cir. 2009) (saying that apportionment could not be reduced to a "mere counting of lines of code," but that the imbalance must "impact the analysis" of how much profit can be property attributable to the infringing feature); Finjan, Inc. v. Sophos, Inc., No. 14-CV-01197-WHO, 2016 WL 4268659, at *3 (N.D. Cal. Aug. 15, 2016) (denying motion to exclude expert testimony counted features to apportion damages).

[26]Ziilabs, *supra* note 25 at *5.

[27]Amended Civil Minutes at 13, California Inst. of Technology v. Broadcom Ltd., No. CV 16-3714-GW(AGRx), (C.D. Cal. Nov. 21, 2019) (No. 1924).

The Central District of California excluded expert testimony on all four of these approaches and went so far as to label counting the number of pages in the 802.11 standard as "lame."[28] Soon thereafter, Caltech was able to obtain one of the largest patent verdicts ever, $1.1 billion.[29]

Another technique that has become increasingly important in high-stakes patent cases is conjoint analysis, a survey methodology that attempts to quantify the value of a given feature by looking at the trade-offs consumers are willing to make.[30] Because it creates a mathematical model using information from a sample of potential customers, conjoint analysis appears to take a more scientific approach to apportionment. The most prominent application of its use in patent litigation is in the smartphone patent wars between Apple and Samsung, where Apple used conjoint analysis to obtain a $1 billion jury verdict against Samsung.[31]

Given conjoint analysis's potential importance in determining patent damages in modern multicomponent products, this article seeks to take a rigorous look at the approach by conducting two surveys and then employing different types of statistical approaches that experts might use to analyze the data.

---

[28]*Id*. at 14. Ironically, one of CalTech's experts had apparently used that approach previously in a different lawsuit. *Id*. at n.15.

[29]Scott Graham, *Caltech and Quinn Emaunel Score $1.1B Verdict Against Apple, Broadcom*, The Recorder (Jan. 29, 2019). Apparently, Apple did not even offer damages testimony after its apportionment arguments were all excluded.

[30]*See* Greg Allenby, Peter E. Rossi, Lisa Cameron, Jeremy Verlinda & Yikang Li, *Calculating Reasonable Royalty Damages Using Conjoint Analysis*, 45 AIPLA Q. J. 233, 234 n.1 (2017) [hereinafter Allenby *et al*., *Calculating Reasonable Royalty Damages*] (listing a number of high-stakes patent litigations where plaintiffs' damages experts used conjoint analysis). *See generally* J. Gregory Sidak & Jeremy O. Skog, *Using Conjoint Analysis to Apportion Patent Damages*, 25 Fed. Cir. B. J. 581 (2016).

[31]*See* Josh Lowensohn, *Jury Awards Apple More Than $1B, Finds Samsung Infringed*, CNET (Aug. 24, 2012, 3:53 PM), https://www.cnet.com/news/jury-awards-apple-more-than-1b-finds-samsung-infringed/. That award was reduced through subsequent litigation and then eventually settled. *See* Scott Graham, *Apple and Samsung Call a Truce in Long-Running Smartphone War,* The Recorder (June 27, 2018), https://www.law.com/legaltechnews/2018/06/27/apple-and-samsung-call-a-truce-in-long-running-smartphone-war/.

II. Conjoint Analysis

Although conjoint analysis is relatively new to the law, it has been widely used in marketing research for years.[32] In a typical conjoint analysis survey, the combination of features for a given product is systematically varied. Respondents react to different product profiles. A computer can then statistically calculate utility scores for each of the separate features (typically using some form of regression analysis). These utility scores are called part-worths.[33]

While there are many forms of conjoint analysis, damages experts typically use Choice-Based Conjoint (CBC) because it approximates the choices consumers face in the real world.[34] In these conjoint surveys, respondents are blinded to the specific feature of interest. Instead, they are asked to choose between different variations of the same product. To simulate real-world choices that consumers face, most surveys also include a "none" choice that can be selected if none of the combinations of features appeal to the survey respondent.[35] Table 1 shows a simple version of the choices that we gave our respondents in the first low-value experiment that we ran.

Of course, modern technology products have countless more features than the five smartphone features found in Table 1. Unfortunately, when a survey presents respondents with too many features,

---

[32]*See* Bryan K. Orme, Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research 5 (4th ed. 2019) ("Today, thousands of conjoint studies are conducted each year over the Internet, via hand-held and mobile technologies, or using person-to-person interviews."); Suneal Bedi & David Reibstein, *Damaged Damages: Errors in Patent and False Advertising Litigation* 12 (Kelley Sch. of Bus., Working Paper No. 19-41, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3440817# ("[C]onjoint analysis is a well-accepted survey methodology that has been used in various business applications.").

[33]*See* Angela Hausman, *Marketing Research: What's Your Part-Worth?*, Market Maven (Nov. 1, 2010), https://www.hausmanmarketingletter.com/marketing-research-whats-your-part-worth/.

[34]Orme, *supra* note 32, at 4; Sidak & Skog, *supra* note 30, at 591–92 (saying that CBC is the proper type of conjoint analysis for reasonable royalty calculations).

[35]Greg M. Allenby, Jeff Brazell, John R. Howell & Peter E. Rossi, *Valuation of Patented Product Features*, 57 J.L. & Econ. 629, 644 (2014) [hereinafter Allenby *et al.*, *Valuation*] (explaining that to be realistic, surveys need to include the none option).

Table 1: Smartphone Choices

Option 1	*Option 2*	*Option 3*	
2-Microphones, some background noise reduction	1-Microphone, no background noise reduction	3-Microphones, superior background noise reduction	None
22 hours battery	20 hours	19 hours	
Service provider can sell your data	Service provider can sell your data	Service provider cannot sell your data	
IP67 (preventing any dust intrusion and resisting water damage at depth up to 1 m for up to 30 mins)	IP55 (resisting some dust and water damage from a low-pressure jet)	IP66 (preventing any dust intrusion and resisting water damage from a high-pressure jet)	
$1098	$798	$949	

respondents tend to resort to simplifying tactics that will distort their true preferences.[36] The upper limit on the number of features that a conjoint analysis survey can present is about six or seven.[37] As a result, respondents see only a subset of the product's features and are told to assume that everything else about the choices are the same.[38]

CBC surveys also contain an important price "feature" as shown in the bottom row of Table 1. By taking the price feature and calculating part-worth utility scores, conjoint analysis can translate utility scores into dollars, which in turn allows researchers to derive respondents' WTP values for each feature in monetary terms. Importantly, the WTP generated by CBC does not measure how much respondents will pay for a particular feature in a vacuum, but instead how much more or less they are willing to pay for one level of a feature rather than a different level of that feature (e.g., how much a respondent will be willing to pay for twenty-two

[36] *See* James R. Betterman, Mary Frances Luce & John W. Payne, *Construction Consumer Choice Processes*, 25 J. Consumer Research. 187, 199 (2008) ("Studies of the effects of problem size show that increases in the number of alternatives facing the consumer lead to greater use of noncompensatory strategies that eliminate alternatives."); Paul E. Green & V. Srinivasan, *Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice*, 54 J. Marketing 3, 8–9 (1990).

[37] Bedi & Reibstein, *supra* note 32, at 16 ("[R]esearch has shown that respondents are overwhelmed with more than six or seven features."); Green & Srinivasan, *supra* note 36, at 8.

[38] *See generally* Allenby *et al., Valuation*, *supra* note 35, at 648.

hours of battery life instead of nineteen hours).[39] This technique is particularly useful in calculating patent damages because the relevant legal question is not how much the patented feature is worth in the abstract, but rather how much the patented feature is worth as compared with the best non-infringing alternative.[40]

The WTP value reflects the entire group of respondents' average preferences, and this average is typically the value used by conjoint analysis practitioners, including expert damages witnesses, to explain how much a particular feature is worth.[41] Importantly, experts have some discretion in selecting among a number of different methodological approaches to calculate WTP values. These choices involve (1) how to trim the data set in an attempt to provide more accuracy; and (2) whether to use average or individual price sensitivities to transform part-worth utilities to dollars.[42] Later in this article, we show how these choices can be manipulated to change WTP results significantly.[43] Even putting aside these problems, courts and commentators have noted a number of other problems regarding the basic idea of using WTP values to set reasonable royalties. We discuss three here.

First, conjoint analysis may not accurately be able to estimate consumer preferences in monetary terms. Brian Orme, the president of Sawtooth Software (the leading purveyor of conjoint analysis software), has cautioned that expressing conjoint utilities in dollar equivalents can be a

---

[39]ORME, *supra* note 32, at 86 (describing "the fallacy of interpreting average dollar values, without respect to competitive offerings."); Visteon Glob. Technologies, Inc. v. Garmin Int'l, Inc., No. 10-CV-10578, 2016 WL 5956325, at *5 (E.D. Mich. Oct. 14, 2016).

[40]There is some dispute over whether the cost of an available non-infringing alternative serves to provide a cap on patent damages or is merely a factor to be considered. William F. Lee & A. Douglas Melamed, *Breaking the Vicious Cycle of Patent Damages*, 101 CORNELL L. REV. 385, 422 (2016); *see also* Thomas F. Cotter, *Four Principles for Calculating Reasonable Royalties in Patent Infringement Litigation*, 27 SANTA CLARA COMPUTER & HIGH TECH. L.J. 725, 743–44 (2011) (explaining why a patent's expected contribution in relation to the next-best alternative is highly relevant to a royalty determination).

[41]Allenby *et al., Valuation*, *supra* note 35, at 630-31 (citing to Apple Electronics Co. Ltd. v. Samsung Elec. No. 11-CV-01846-LHK, 2001 WL 7036077 (N.D. Cal. Dec. 2, 2011) and Microsoft Corp. v. Motorola Inc., 696 F.3d 872 (9th Cir. 2012)).

[42]*See infra* notes 85–88 and accompanying text.

[43]*See infra* Part IV(B) & (C).

trap.[44] He notes that "[e]ven when computed reasonably, the results often seem to defy commonly held beliefs about prices and have limited strategic value for decision making."[45] Orme does not point to any study to support his claim, and we suspect his understanding is based on his experience working in the private sector. There is at least some available data to assess. In *Apple v. Samsung*, Apple's expert, MIT's John Hauser, testified that his conjoint analysis survey showed that consumers were willing to pay $39 more to have the so-called reverse pinch zoom feature in a $199 smartphone.[46] It is hard to see how this "tiny tweak" increased the value of a smartphone by 15 percent.[47] Unfortunately, we cannot fully evaluate the work done in that case because the complete analysis is not available. More recently, surveys by Suneal Bedi and David Reibstein (the latter served as Samsung's expert in its lawsuit with Apple) produced absurdly high numbers. Their surveys supposedly found that consumers are willing to pay anywhere from $6,815 to as much as $10,161 to have four cup holders in their car instead of one.[48]

A second critique of conjoint analysis is that a consumer's willingness to pay does not determine what the product's market price will be. In other words, simply because an average consumer may be willing to pay $25 more for an infringing feature does not mean that a company can sell its product for $25 more. When products are sold in the real world, there are a number of other factors, such as competition and profitability, that need to be taken into account.[49]

---

[44]ORME, *supra* note 32, at 85 (labeling the section that discusses the expression of conjoint utilities in monetary terms as "Monetary Scaling Trap").

[45]ORME, *supra* note 32, at 85.

[46]Bernard Chao & Roderick O'Dorisio, *Saliency, Anchors & Frames: A Multicomponent Damages Experiment*, 26 MICH. TECH. L. REV. 1, 11–12 (2019) (describing Hauser's testimony and reproducing the demonstrative exhibit he used).

[47]*See* Joe Mullin, *Snapping Back: Appeals Court Revives Apple's Patented "Rubber Banding" Tech Because of One Small Tweak*, ARSTECHNICA (Apr. 17, 2017, 10:22 AM) (explaining how minor an invention Apple's patent covers).

[48]Bedi & Reibstein, *supra* note 32, at 48.

[49]Visteon Glob. Technologies, Inc. v. Garmin Int'l, Inc., No. 10-CV-10578, 2016 WL 5956325, at *6 (E.D. Mich. Oct. 14, 2016) (noting with approval an expert report that explained that actual price is determined by consumer value, produce costs and competition, but that conjoint analysis reflects only consumer value); ORME, *supra* note 32, at 86 (using conjoint analysis to place a monetary value on a feature "assumes no competition."); Allenby *et al., Valuation*, *supra* note 35, at 661 ("Incremental profits are determined not only by the value that individual consumers place on the patented feature but also on the nature and extent of competition in the relevant market.").

This reality affects the hypothetical negotiation because the infringer probably does not have $25 more per unit to split with the patent holder.

To solve this problem, Greg Allenby and his colleagues suggest calculating what they call "equilibrium outcomes" to determine the "incremental economic profits that would accrue to a firm as a product is enhanced by patented features as a way of computing the economic value of a patent."[50] This approach still relies on conjoint analysis, but instead of using consumers' average preferences to produce a per unit dollar value of a feature, this approach examines the range of consumers' preferences.[51] It specifically looks at individual preferences to determine what portion of consumers would still have bought the defendant's "defeatured" product.[52] It then calculates how profits differ in the infringing world from the non-infringing world.[53] To date, this approach does not appear to have been used in litigation.[54] Perhaps it is because, as currently framed, the technique assumes that the patentee and infringer are competitors. In reality, reasonable royalty calculations are often used precisely because the patentee does not compete with the infringer.[55] We suspect that small adjustments to Allenby and his colleagues' methodology would make it suitable for calculating royalties for the more common, nonpracticing entity plaintiff, assuming that conjoint analysis produces accurate demand side valuations in the first place.

Finally, a third potential problem with conjoint analysis is that its surveys are subject to manipulation. For example, Bedi and Reibstein show

---

[50]Allenby *et al., Valuation*, *supra* note 35, at 631.

[51]When we use the term "conjoint analysis," we are referring to choice based conjoint since that is the only flavor of conjoint analysis used in patent litigation.

[52]Allenby *et al.*, *Calculating Reasonable Royalty Damages*, *supra* note 30, at 247–48.

[53]Allenby *et al.*, *Calculating Reasonable Royalty Damages*, *supra* note 30, at 248. This analysis also requires knowledge of the cost of producing both the infringing and non-infringing products to determine profits.

[54]There are two variations to Allenby's approach, the "market share method" and the "equilibrium profit method." Allenby *et al.*, *Calculating Reasonable Royalty Damages*, *supra* note 30, at 234–35. Westlaw searches conducted on May 11, 2020, using both (1) "market share method" and "conjoint analysis" and (2) "equilibrium profit method" and "conjoint analysis" yielded no briefs and no decisions.

[55]Mark A. Lemley, *Distinguishing Lost Profits from Reasonable Royalties*, 51 Wm. & Mary L. Rev. 655, 661 (2009) ("Reasonable royalty law is designed with the nonmanufacturing patentee in mind.").

that the results of a conjoint analysis survey can be manipulated by including minor features instead of major features in the experimental design.[56] They conducted two surveys that determined how much respondents valued both cup holders (one, two, or four cup holders), and gas cap location in a car (passenger side versus driver side). One survey included three "important" features (brand, miles per gallon, and vehicle type) with the two features of interest.[57] The other survey included "unimportant" features (clock style, handle type, and coin slot) with the two features of interest.[58] They found that the monetary value of both the number of cup holders and the location of the gas cap increased dramatically when they ran a CBC survey that included only unimportant features. Bedi and Reibstein theorize that when only unimportant features are included, the survey is artificially focusing the respondents' attention on the features of interest and causing them to "overweight" these features.[59]

In sum, there are several serious questions regarding the use of conjoint analysis in patent cases. First, there are some indications that conjoint analysis might be inaccurate and provide unrealistically high valuations. Second, it is unclear whether willingness to pay is the correct measurement because conjoint analysis (in its current incarnation) seeks only to examine the demand side of a patented feature. Third, experts might be able to manipulate outcomes by intentionally including low-value "other features" in the study design. Importantly, even if these issues create real

---

[56]Bedi & Reibstein, *supra* note 32, at 48 and 49.

[57]The "Brand" was either a Ford, Toyota, or Volkswagen. Miles per gallon was either 30, 40, or 50. Vehicle type was either Coupe, Sedan, or SUV. Bedi & Reibstein, *supra* note 32, at 47 tbl. 2.

[58]Clock style was analog, digital, or both. Handle type was flat, bottom, or top. Coin slot was none, one, or two. Bedi & Reibstein, *supra* note 32, at 47 tbl. 3.

[59]Bedi & Reibstein, *supra* note 32, at 31–35 (suggesting that the problem can be further broken down into three issues that the authors call "binary choice, focalism and lack of awareness"); *see also* Joel Huber, *What We Have Learned from 20 Years of Conjoint Research: When to Use Self-Explicated, Graded Pairs, Full Profiles or Choice Experiments*, Sawtooth Software: Research Paper Series, 2 (1997), https://sawtoothsoftware.com/resources/technical-papers/what-we-have-learned-from-20-years-of-conjoint-research-when-to-use-self-explicated-graded-pairs-full-profiles-or-choice-experiments ("Evaluation tasks intentionally force respondents to attend to attributes that they might otherwise not notice. In doing so, attention can elevate the importance of particular attributes to a level that is greater than would occur in the marketplace.").

problems, it does not mean that conjoint analysis should never be used in a legal context. Orme for instance argues that conjoint analysis is still useful to show consumers' relative preference among different features.[60]

This article examines the second and third issues discussed above. Does conjoint analysis yield reasonably accurate values? Can experts manipulate outcomes to favor their clients? For the latter question, we go beyond what Bedi and Reibstein studied and seek to assess whether conjoint analysis is susceptible to different types of "data fishing."

III. The Experiments

We sought to replicate the work patent damages experts are called to perform by conducting our own conjoint analysis experiments. Our goals were to determine whether conjoint analysis produced reliable results and whether the experiments could be manipulated to favor one side or the other. As part of our experimental design, we varied the "other" features that were included in two surveys to determine whether they would affect the prices of the features of interest as Bedi and Reibstein's survey results did. Finally, we were also concerned that experts can manipulate the results of a conjoint analysis survey by choosing which statistical approach produces the best result for their client. Accordingly, we applied a variety of statistical approaches to our data to expose the dangers of allowing an expert to simply select the one that favors her client.

Our surveys used conjoint analysis to assess the value of particular smartphone features. Smartphones contain thousands if not hundreds of thousands of patented features.[61] Moreover, there have been so many patent lawsuits involving smartphones that they have collectively been referred to as the "smartphone wars."[62] We selected six different smartphone

[60]Orme, *supra* note 32, at 26 ("[C]onjoint models are excellent directional indicators.").

[61]In 2011, RPX estimated that 250,000 patents are relevant to the smartphone. RPX Corp., Registration Statement (Form S-1) (Sept. 2, 2011), https://www.sec.gov/Archives/edgar/data/1509432/000119312511012087/ds1.htm. Since 2011, Smartphones have become even more complicated and are likely covered by even more patents every year.

[62]Rajdeep Banerjee & Joyeeta Banerjee, *Smartphone Patent Wars: No End in Sight*, Forbes India (June 1, 2018, 3:30 PM), https://www.forbesindia.com/blog/technology/smartphone-patent-wars-no-end-in-sight/ (discussing multiple patent lawsuits involving Apple, Samsung, Qualcomm, HTC, Ericsson, Sony, Motorola, and Microsoft, among others).

features and conducted a pretest to determine which features were more valuable and which features were less valuable. We then constructed two conjoint surveys that included our two features of interest and either two additional low-value features or two additional high-value features.

A. Feature Selection

We chose features regularly described in smartphone advertisements along with what we hoped customers would assign as a range of perceived values. The selected smartphone hardware features included background noise reduction (via multiple microphones), battery life, screen resolution, storage (solid-state drive size), and water resistance. Further, we included a data privacy option, a feature prohibiting the mobile carrier from selling a user's location data. Researchers have found that privacy has been notoriously difficult to value.[63] Finally, we varied price as the final "feature." For each feature, we chose between two and four different levels. We chose feature levels that could be explained in short, digestible descriptions. Table 2 shows the different features and associated levels.

B. Pretest

After selecting features that spanned a spectrum of perceived value, we created short descriptions of each feature and icons to visually represent each feature level. We designed the descriptions and icons to convey information in an easily digestible way so that survey takers could make quick but informed decisions. Figure 1 illustrates both the description and icons for the battery life feature. Similar descriptions and icons were created for each feature found in our surveys.

After creating smartphone feature descriptions and icons, we needed to assess how people viewed their relative value. Thus, we created a short survey that asked participants to rank the six described features in order of value. We conducted our pretest survey using the MTurk

[63]*See* Angela G. Winegar & Cass R. Sunstein, *How Much Is Data Privacy Worth? A Preliminary Investigation*, 42 J. Consumer Pol'y 425 (2019) (finding a large discrepancy in individuals' willingness to pay (WTP) for privacy as compared with their willingness to accept (WTA) payment); Kirsten Martin & Helen Nissenbaum, *Measuring Privacy: An Empirical Test Using Context to Expose Confounding Variables*, 18 Colum. Sci. & Tech. L. Rev. 176, 218 (2016) (arguing that the difference between reported privacy preferences and privacy behavior is caused by the lack of understanding of privacy practices and suggesting that many survey instruments do not adequately address this issue).

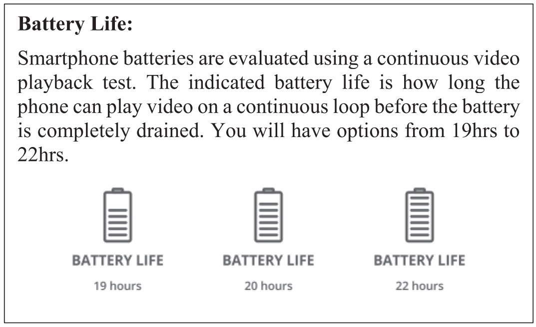
Table 2: Feature Levels

Background Noise Reduction	1-Microphone, no background noise reduction
	2-Microphones, some background noise reduction
	3-Microphones, superior background noise reduction
Battery Life	19 hours
	20 hours
	22 hours
Privacy	Service provider can sell your location data
	Service provider cannot sell your location data
Screen Resolution	720p (pre-HD)
	1080p (HD)
	2160p (4K)
Storage	128 GB
	256 GB
	512 GB
	1 TB
Water Resistance	IP55 (resisting some dust and water damage from a low-pressure jet)
	IP66 (preventing any dust intrusion and resisting water damage from a high-pressure jet)
	IP67 (preventing any dust intrusion and resisting water damage at depth up to 1 m for up to 30 mins)
	IP68 (preventing any dust intrusion and resisting water damage at depths over 3 m for longer than 30 mins)
Price	$649
	$798
	$949
	$1098

platform.[64] We paid $1.50 per survey to a total of 109 pretest participants. We defined value for the participants as "which feature's best option is the most valuable as compared to the less desirable option or options within the same group." Survey participants evaluated the six features listed in Table 2 (excluding price). Participants assigned a rank of one to the feature with the most perceived value and a rank of six to the feature with the least perceived value. We used the mean results of the pretest to inform our conjoint survey designs. The results of our pretest are shown in Figure 2 below. The results appear more or less consistent with intuition. Respondents valued battery life the most and water resistance the least.

---

[64]*See* Krin Irvine, David A. Hoffman & Tess Wilkinson-Ryan, *Law and Psychology Grows Up, Goes Online, and Replicates*, 15 J. Empirical Legal Stud. 320, 344 (2018) (discussing the use of Amazon MTurk's crowdsourcing platform and explaining why it provides an appropriate subject pool).


Battery Life:

Smartphone batteries are evaluated using a continuous video playback test. The indicated battery life is how long the phone can play video on a continuous loop before the battery is completely drained. You will have options from 19hrs to 22hrs.

BATTERY LIFE 19 hours

BATTERY LIFE 20 hours

BATTERY LIFE 22 hours

Figure 1. Battery Life Feature Description

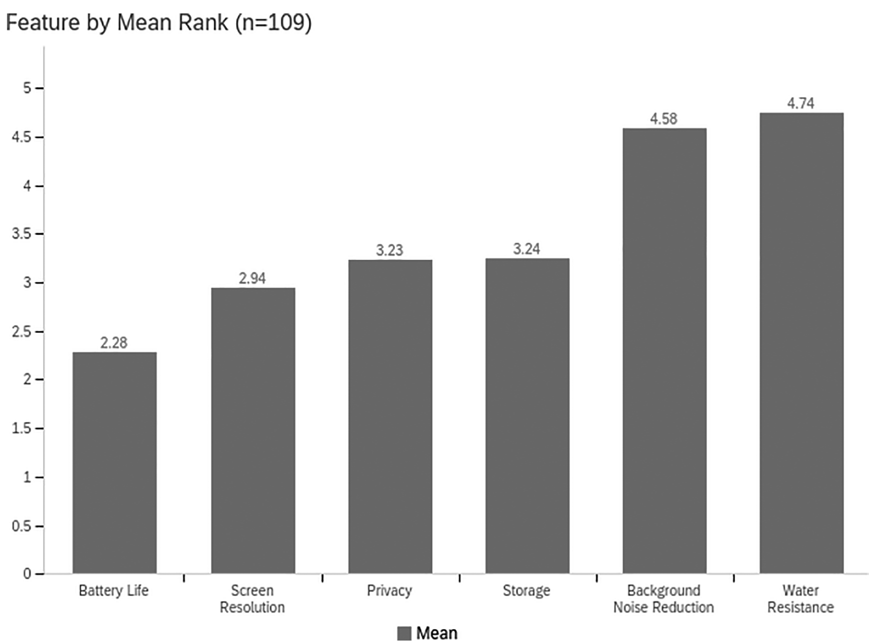


Figure 2. Preliminary Valuation Survey Ranking Results

C. The Two Conjoint Surveys

We tested two common features across our two surveys: (1) the most valuable feature, battery life; and (2) location privacy. As shown in Table 3 below, in the first survey, “Low-value Conjoint,” we tested the two features of interest with the least valuable features: background noise

Table 3: Features Used in Experiments

Feature (# of levels)	*Low-value conjoint*	*High-value conjoint*
Screen resolution (3)		X
Storage (4)		X
Battery life (3)	X	X
Location privacy (2)	X	X
Background noise reduction (3)	X	
Water resistance (4)	X	
Price (4)	X	X

reduction and water resistance. In the second survey, "High-value Conjoint," we tested the same two features against the most valuable remaining features: screen resolution and storage. Both experiments also contained four price levels of the smartphone ($649, $798, $949, and $1098). Including price allows us to assess different trade-offs in terms of dollars. Table 3 summarizes each survey set.

For both the low-value and high-value conjoint surveys, we configured the survey to allow participants to choose between purchasing three different product packages or to not purchase any package. Figure 3 shows one possible choice that a Condition B survey participant may have seen. In the actual surveys, participants saw three such choices.

Each survey included demographic questions as well as simple "attention check" questions and timing metadata. We included demographic questions for informational purposes, but we used the attention checks and timing metadata to filter out unreliable participants. We included a series of three randomized attention checks as an early filter. We terminated surveys and did not pay participants who failed the randomized attention checks before answering the conjoint survey. We also included a final attention check following the conjoint analysis. The question asked participants to identify which feature was not part of the conjoint survey. We also terminated surveys of and did not pay participants who failed this question. Further, we tracked how long participants spent reading the instruction and feature description pages. In the final analysis, we did not include participants who spent less than two seconds reviewing any of the instruction or feature description pages. We paid $3 per survey for both Low-value and High-value Conjoint. Low-value Conjoint was run at 3:00 p.m. on January 10, 2020, and excluded any participant who had taken the pretest. High-value Conjoint was run precisely one week later also excluding anyone who had participated in the pretest or

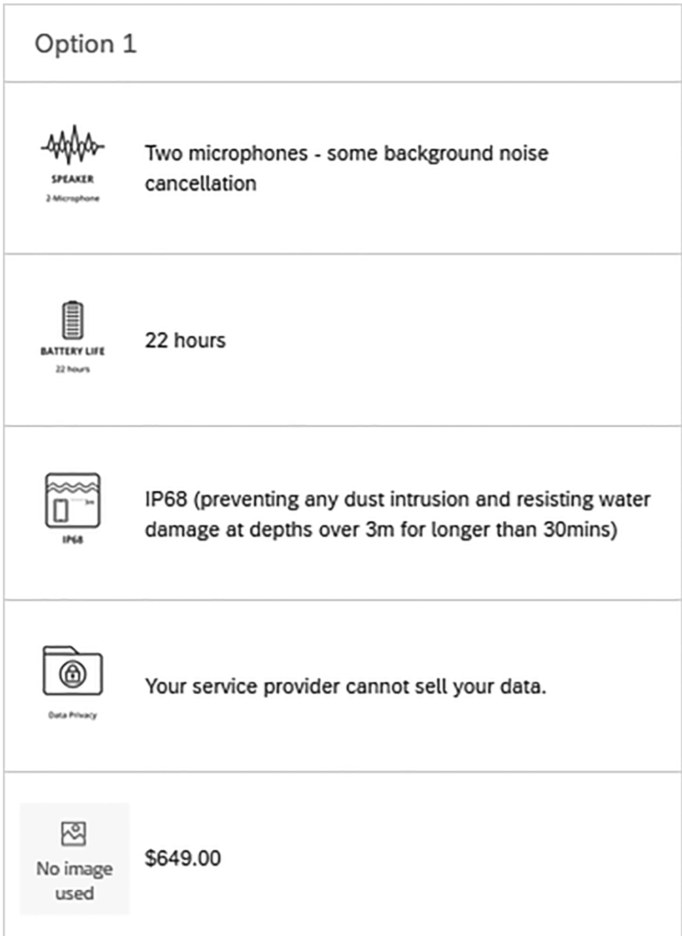


Figure 3. Example, Options from Low-value Conjoint Survey

Low-value Conjoint survey. Participants located outside the United States and under the age of eighteen were also excluded from each survey.

1. Low-Value Conjoint

Five hundred two people participated in the Low-value Conjoint survey. Thirty-six participants (7.2 percent) failed the initial randomized attention checks. Another seventy-four participants (14.7 percent) failed the final survey specific attention check, leaving us with 392 paid participants.

Of those, we filtered out an additional twenty-nine participants (5.8 percent of those who started the survey, or 7.4 percent of the paid participants) who clicked through any of the instructions or feature description pages in less than two seconds. After filtering, we included 363 participants in the final analysis. Of the 363 participants, a majority were male, the average age was thirty-seven, 75 percent were white, and most had at least a bachelor's degree.[65] The demographics of our survey also leaned more educated and liberal than the general population.[66]

2. High-Value Conjoint

Four hundred eighty-two people participated in the High-value Conjoint survey. Thirty-five participants (7.3 percent) failed the initial randomized attention checks. Another forty-one participants (8.5 percent) failed the final survey specific attention check, leaving us with 406 paid participants. Of those, we filtered out an additional sixty-one participants (12.7 percent of those who started the survey, or 15.0 percent of the paid participants) who clicked through any of the instructions or feature description pages in less than two seconds. After filtering, we included 345 participants in the final analysis. Of the 345 participants, a majority were male, the average age was thirty-seven and one-half, 75 percent were white, and most had at least a bachelor's degree.[67] The

[65]Sex: 59.5 percent Male, 40 percent Female, 0.5 percent Neither; Age: Average of 36.8, Standard Deviation of 11.3; Ethnicity: 74.9 percent White, 14.1 percent Black, 0.8 percent American Indian, 6.9 percent Asian, 0.3 percent Islander, 3 percent Other; Hispanic: 88.9 percent No, 5 percent Mexican / Mexican American / Chicano, 1.1 percent Puerto Rican, 0.6 percent Cuban, 4.4 percent Other Latino.

[66]Highest Level of Education: 0.3 percent Doctorate, 0.8 percent Professional Degree, 9.4 percent Master's, 42.4 percent Bachelor's, 11.3 percent Associate's, 20.1 percent Some College Credit, 14.9 percent High School (Diploma or GED), 0.8 percent Some High School; Household Income: 0 percent $200,000 or more, 2.2 percent $150,000 to $199,999, 7.4 percent $100,000 to $149,999, 37.5 percent $50,000 to $99,999, 28.9 percent $30,000 to $49,999, 18.7 percent $10,000 to $29,999, 5.2 percent Less than $10,000; Political Affiliation Scale: 26 percent Strong Democrat, 18.2 percent Democrat, 12.2 percent Slight Democrat, 14.1 percent No Preference, 9.4 percent Slight Republican, 9.9 percent Republican, 10.2 percent Strong Republican.

[67]Sex: 53.3 percent Male, 45.5 percent Female, 1.2 percent Neither; Age: Average of 37.5, Standard Deviation of 11.9; Ethnicity: 76.8 percent White, 13.9 percent Black, 1.2 percent American Indian, 6.4 percent Asian, 0.3 percent Islander, 1.45 percent Other; Hispanic: 91.2 percent No, 3.2 percent Mexican / Mexican American / Chicano, 1.5 percent Puerto Rican, 0 percent Cuban, 4.1 percent Other Latino.

demographics of our survey leaned more educated and liberal than the general population.[68]

IV. The Results

Qualtrics provided individualized participant utility scores for each feature level and price level.[69] Using this information, the platform automatically calculates prices for each feature. However, it does not allow researchers to choose among different statistical approaches. Because we wished to study how experts might work with the underlying data, we took the individual utility scores and applied a variety of statistical approaches to the data to see if these choices significantly affected the results that an expert would report. Such a finding would suggest that conjoint analysis may not be sufficiently robust to be of much practical use. Additionally, it would show how experts might manipulate their analyses to favor the party that they represent.

For the purpose of Section A, we simply report the results from the approach that we believe conveyed the data most accurately. Specifically, we trimmed the data set by removing 5 percent of the responses with the highest price sensitivity and 5 percent with the lowest price sensitivity. We then used each respondent's individual price sensitivity to calculate their willingness to pay. The WTP values we report are disproportionately high given the number of features found in a smartphone. Moreover, although our results provided mostly unrealistically high valuations, that outcome was not uniform. As a result, we could not

[68]Highest Level of Education: 0 percent Doctorate, 1.5 percent Professional Degree, 13 percent Master's, 37.7 percent Bachelor's, 12.5 percent Associate's, 24.1 percent Some College Credit, 11 percent High School (Diploma or GED), 0.3 percent Some High School; Household Income: 1.5 percent $200,000 or more, 2.6 percent $150,000 to $199,999, 7.8 percent $100,000 to $149,999, 38 percent $50,000 to $99,999, 22.9 percent $30,000 to $49,999, 22.3 percent $10,000 to $29,999, 4.9 percent Less than $10,000; Political Affiliation Scale: 22.6 percent Strong Democrat, 17.1 percent Democrat, 14.2 percent Slight Democrat, 14.8 percent No Preference, 13 percent Slight Republican, 9.6 percent Republican, 8.7 percent Strong Republican.

[69]Qualtrics uses Hierarchical Bayes estimation to derive "individual-based utility models." *Conjoint Analysis White Paper*, Qualtrics, https://www.qualtrics.com/support/conjoint-project/getting-started-conjoints/getting-started-choice-based/conjoint-analysis-white-paper/ (last visited Dec. 28, 2020). Qualtrics does not provide the underlying individual choices each respondent made.

merely scale down the results to arrive at more realistic values. Importantly, the statistical choices we made did not affect the findings in section A. All approaches yielded disproportionately high results.

However, one tactic had a huge impact on the results' variance. In section B, we explain how using individual price sensitivity (as opposed to average price sensitivity) dramatically changed the size of the confidence interval an expert could report. Tight confidence intervals generally suggest that results are reliable. Here, we believe the use of average price artificially removes variability from the results, giving readers a false sense of reliability.

In section C, we also examine different techniques for filtering data. While all these approaches resulted in extremely high values, the differences among the approaches were certainly large enough to matter to parties in litigation. Thus, there is a danger that experts will engage in "data fishing," the practice of choosing an approach that intentionally favors one party.[70] Finally, in section D, we report on efforts to show that results can be manipulated by changing the "other" features tested in the conjoint analysis survey. Unlike Bedi and Reibstein, we did not find that altering the feature mix had a significant impact on the value of our two control features.

A. Disproportionately High Values

Table 4 depicts the results from our two surveys.[71] The WTP values reflect the amount participants were willing to pay in order to obtain an upgraded feature instead of the baseline feature. For example, if we consider "Battery Life," our results suggest that the average consumer was willing to pay $117.73 to obtain a smartphone with a battery that lasted twenty hours (the upgraded feature) instead of nineteen hours (the

---

[70]Hillel J. Bavli, *Credibility in Empirical Legal Analysis* 3 (SMU Dedman Sch. of L. Legal Stud., Research Paper No. 434, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3434095 ("Data fishing, also known as data dredging or p-hacking, is a well-recognized problem in the hard and social sciences that involves using data to search for and selectively (and misleadingly) report results that are statistically significant or otherwise favorable to the researcher.").

[71]As mentioned earlier, the results in Table 5 were calculated by trimming 5 percent of the top and 5 percent of bottom respondents based on price sensitivity, and then using individual price sensitivity to calculate each respondent's WTP. The confidence intervals for these findings are found in Table A of the Appendix. We discuss issues with these confidence intervals later.

Table 4: Willingness to Pay Results

Feature	*Baseline*	*Upgraded*	*WTP Low-value conjoint*	*WTP High-value conjoint*
Location privacy	Location sold privacy	Not sold	$569.11	$557.97
Background noise reduction	One Microphone	Two microphones	$320.91	
		Three microphones	$470.68	
Battery life	19 hours	20 hours	$111.73	$92.35
		22 hours	$186.54	$135.63
Water resistance	IP55	IP66	$151.35	
		IP67	$206.54	
		IP68	$241.86	
Screen resolution	720p (pre-HD)	1080p (Full HD)		$295.90
		2160p (4K)		$603.97
Storage	128 GB	256 GB		$51.61
		512 GB		$132.40
		1 T		$220.66

Table 5: Difference between Low- and High-value Conjoints

	Battery (20 hours)	*Battery (22 hours)*	*Location Privacy*
Mean	$19.38	$50.91	$11.15
2.5%–97.5%	$4.75–$34.54	$26.25–$76.62	$(83.89)–$108.06

baseline feature). Both surveys included the "Location Privacy" and "Battery Life" features. Accordingly, those features have two WTP values reflecting the results from the different surveys. Otherwise, each feature only has one WTP value.

As a group, the WTP values were unrealistically high. They ranged from $51.61 for 256 GB to $603.97 for 2160p (4K) screen resolution. Standing in isolation, these values may not appear extreme. However, if we consider the number of patents that cover a smartphone, the problem becomes apparent. Recall, in 2011, RPX estimated that there were roughly 250,000 patented features in a smartphone.[72] This estimate seems reasonable (at least as to the order of magnitude) given that

[72]*See supra* note 61.

smartphones use many different standards (e.g., Wi-Fi, Bluetooth) and each of these standards can be covered by hundreds to tens of thousands of patents.[73] If anything, the number of patented features is growing as the technology becomes more complex.

We can estimate the average consumer's WTP for all the features in a smartphone assuming that there are *only* 100,000 patented features in a smartphone (not 250,000) and using the *lowest value* our surveys yielded ($51.61) as the average value for a patented feature. Using these conservative baselines, our study suggests that consumers would be willing to pay roughly $5,161,000 for the 100,000 patented features. Of course, it is absurd to suggest that any consumer, let alone the average consumer, would pay more than $5 million for a feature-rich smartphone. Perhaps all the features found in our survey were far more valuable than the average feature, but even if we cut our lowest price by a *factor of one hundred*, we end up with a WTP of more than $50,000 for a smartphone.

Another way to interpret the results is to ignore all the thousands of features that are not found in our surveys. We simply compare the WTP for a smartphone with all the best features against a smartphone with the corresponding basic features. Our results suggest that an average consumer would pay $2241.91 more for a smartphone with location privacy ($569.11), three microphones ($470.68), twenty-two hours of battery life ($135.63), IP68 water resistance ($241.86), 2160p (4K) screen resolution ($603.97), and 1T of storage ($220.66) than for a smartphone with the baseline features.[74] While this result does not seem quite as absurd, it is unrealistic to suggest that an average consumer would pay more than $2200 more for this smartphone given that the most expensive smartphones today are around $1500.[75] Keep in mind this number ignores thousands of other patented features.

---

[73]*See* Jorge L. Contreras, *Technical Standards, Standards-Setting Organizations and Intellectual Property: A Survey of the Literature (With an Emphasis on Empirical Approaches)*, *in* 2 Research Handbooks On Economics Of Intellectual Property Law 185, 191–93 (2019).

[74]To be conservative, when a feature's value was assessed by both surveys, we used the lower value.

[75]Antonio Villas-Boas, *The Most Expensive iPhone Apple Sells Costs $1,450—About the Same as a 13-inch MacBook Pro*, Business Insider (Sept. 12, 2019, 7:06 AM), https://www.businessinsider.com/apple-iphone-11-pro-max-512gb-costs-same-macbook-pro-2019-9?utm_source=copy-link&utm_medium=referral&utm_content=topbar.

To be clear, we are not the first to suggest that conjoint analysis produces unrealistically high values for features. Orme noted that when conjoint analysis is used to arrive at monetary amounts, the results often reveal "that respondents are willing to pay much more for one feature over another than is suggested by market prices."[76] Bedi and Reibstein's surveys illustrate the same point, and inflated results were also introduced in the *Apple v. Samsung* smartphone wars.[77]

One might think that these numbers could be adjusted downward to achieve more realistic values. The problem with this approach is that conjoint analysis does not inflate values uniformly. Thus, it is unclear what scale could be used. Our results for location privacy and storage illustrate this split. To determine how to adjust the results downward, we would need to find a reference point. Our first attempt uses location privacy. In the real world, consumers are willing to pay very little for privacy. Yet many surveys report that people place a high value on privacy. Many scholars explain this so-called privacy paradox by suggesting that a combination of lack of information, cognitive biases, and context cause consumers to behave in ways that do not reflect how much they truly value privacy.[78] Our study is not intended to contribute to this discussion, but the existence of the discussion shows how hard it can be to find an uncontroversial reference point.

We chose to use \$1.19 as a reference point for the value of location privacy based on a study from Scott Savage and Donald Waldman.[79] Importantly, their study measured one of the features we tested—what consumers were willing to pay for location privacy. Moreover, the authors attempted to account for hypothetical bias by conducting a survey where they assured their respondents that "the apps are real, are traded in markets, and that [the respondents] will be making (or not

---

[76]*See* Orme, *supra* note 32, at 86.

[77]*See supra* text accompanying notes 47 and 48.

[78]*See* Kirsten Martin & Helen Nissenbaum, *What Is It about Location?*, 35 Berkeley Tech. L.J. 251, 278 n.90 (2020); *see also* Kenneth A. Bamberger, Serge Egelman, Catherine Han Elazari Bar On & Irwin Reyes, *Can You Pay for Privacy? Consumer Expectation and the Behavior of Free and Paid Apps*, 35 Berkeley Tech. L. J. 327, 339 (2020).

[79]Scott J. Savage & Donald M. Waldman, *Privacy Tradeoffs in Smartphone Applications*, 137 Econ. Letters 171, 173 (2015).

making) an actual purchase."[80] While it is easy to quibble with this choice, our point is merely to show that conjoint analysis does not inflate values uniformly. If we use $1.19 as the correct price for location privacy, it would suggest that our conjoint analysis survey inflated values by a factor of 478.[81]

Contrast this result with using storage as a reference point. Table 4 suggests that consumers would pay $51.61 to go from 128 to 256 GB. However, in the real world, Apple charges $50 for each additional 64 GB of storage in otherwise identical phones.[82] In other words, conjoint analysis produced prices for storage that were roughly 50 percent less than real-world pricing. If we assume that the Apple price is the correct price for storage, it would suggest that our conjoint analysis deflated values by a factor of two.[83]

In short, for the most part, our two conjoint analysis surveys yielded unrealistically high results. Moreover, we were unable to find a suitable scale to adjust our conjoint analysis results. Using existing reference points, we found that the results were between 478 times too high or two times too low.

B. Average versus Individual Price Sensitivity

The results reported in the previous section were calculated using the methods that we believe most accurately reflect consumer preferences. However, we made two important choices in performing our calculations. First, we chose to use individual price sensitivities instead of average price sensitivities to calculate WTP values. Second, we chose to trim our responses by dropping 10 percent of the responses that provided the most extreme results (5 percent on each end). In this section, we

---

[80]*Id*. at 172.

[81]This number was calculated by simply dividing $569.11, the WTP from our first experiment, by 1.19.

[82]*iPhone*, APPLE, https://www.apple.com/iphone/ (last visited Jul. 5, 2020).

[83]While our pretest was intended only to give a sense of relative valuations, we also note that the results of our conjoint analysis surveys were not consistent with the rankings revealed in our pretest. Our pretest suggested that battery life was the most valuable feature. Yet, our conjoint analysis surveys suggested that location privacy, screen resolution, and background noise were more valuable than battery life.

justify using individual price sensitivities. Indeed, we argue that the failure to use this approach will often result in a misleading level of certainty.

Price sensitivity refers to how sensitive individuals are to changes in price. Different individuals have different price sensitivities, and these price sensitivities can change from one product to another.[84] Conjoint analysis determines an individual's price sensitivity by looking at the difference in part-worth utilities for the different price levels of a given product. For example, in our experiment, the smartphones were priced at four levels. An individual's price sensitivity is calculated by taking the difference between the highest price and lowest price options and dividing the resulting value by the difference in the prices' respective utility scores. For example, the following calculation shows one of our respondent's price sensitivity (\$1098–\$649) / (−1.395–1.265) = −\$168.84/util.[85] The average price sensitivity is calculated by running the same calculation once, using only the average of all participants' individual utility scores for the respective prices.

At this point, an expert has a choice. She can either (1) use a respondent's *individual* price sensitivity to determine each respondent's WTP for each,[86] or (2) take all the respondents' *average* price sensitivity to determine each respondent's WTP. Naturally, individual price sensitivity more accurately reflects individual choices. Nonetheless, it is common practice to use average price sensitivity to calculate individual WTP.[87] For example, we used the Qualtrics conjoint analysis software platform in our experiments. While Qualtrics allowed us to calculate individual price sensitivities using individual utilities, the reports Qualtrics generated were based on average price sensitivities. We had to perform our own calculations to derive WTP values based on individual price sensitivities.

---

[84]Orme, *supra* note 32, at 84.

[85]A negative utility value for price shows that higher prices are undesirable.

[86]*See, e.g.*, Visteon Glob. Technologies, Inc. v. Garmin Int'l, Inc., No. 10-CV-10578, 2016 WL 5956325, at *5 (E.D. Mich. Oct. 14, 2016) ("Dr. Steckel used the respondent-level partworths on price [individual price sensitivity] and the other attributes to estimate the EV [economic value] of the patented features … .").

[87]This appears to be how Qualtrics calculates WTP; *see also* Bedi & Reibstein, *supra* note 32, at 46.

Experts may think that using average price sensitivity to calculate WTP is preferable because it eliminates the problem of outliers. In most conjoint analysis experiments, there are some number of outliers that can skew the results. For example, some individual responses reflect a positive price sensitivity that unrealistically suggests that individuals prefer paying more money for inferior features.[88] To address the outlier problem, experts may choose from several different potential approaches: (1) using average price sensitivities that essentially mask the outlier problem, (2) filtering the data by dropping some outliers, or (3) doing both.[89]

We argue that taking the first approach is a serious mistake. Using average price sensitivities artificially suggests that the range of respondents' WTP is more similar than it actually is. More concretely, when various statistical tests are applied to the resulting WTP data, the results misleadingly show that the conclusions have a high degree of reliability because the resulting WTP values are grouped more closely together than they would have been if we had used individual respondents' actual price sensitivities.

The data from our experiments illustrates this problem. We calculated WTP values for all the features of interest using both average and individual price sensitivities.[90] A comparison of these two approaches yields some differences in WTP values themselves. While these differences would undoubtedly be important to parties in litigation, the results at least had the same order of magnitude. In contrast, when 95 percent confidence intervals of the WTP values were calculated, the size of the confidence intervals differed by an order of magnitude, sometimes

---

[88]*See infra* note 88 and accompanying text (discussing the problem of respondents that exhibit positive price sensitivity).

[89]*See infra* § IV(C) (discussing different possible techniques for filtering data).

[90]Unlike the results from § IV(A), we did not trim the data. However, we did discard one extreme outlier where the respondent's price sensitivity was three orders of magnitude above the median and exclusive mean (mean discarding the outlier). The effect on the "average" price sensitivity calculation was a reduction in the Dollar Value of Utility by less than 50 cents and, on average, a 95 percent reduction of the already small 95 percent confidence interval for each feature level. The result on the individual price sensitivity was, on average, a 79 percent reduction in the 95 percent confidence interval for each feature level. Without the elimination of the extreme outlier, every feature level would have a 95 percent confidence interval (as calculated using individual price sensitivities) that included negative values (i.e., the results would be meaningless).

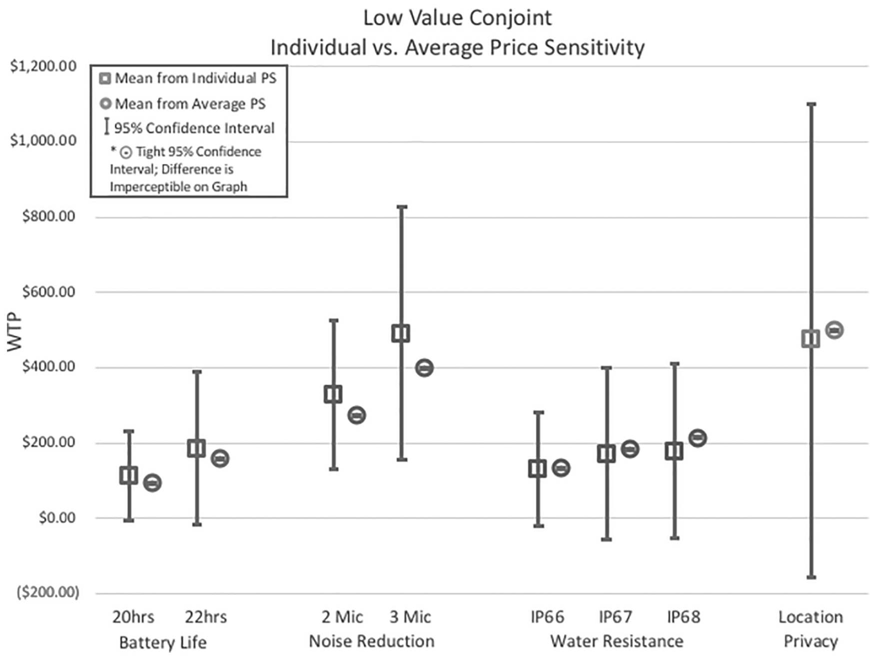


Figure 4. Low Value Conjoint

several. This difference can be observed in Figure 4. Figure 4 compares the average WTP and the 95 percent confidence intervals for each feature in experiment A using individual and average price sensitivity.[91] The 95 percent confidence intervals calculated using individual price sensitivities are visible as vertical bars. The 95 percent confidence intervals calculated using average price sensitivity are so small that they barely appear as visible horizontals lines (underneath each circle).

Consider the results for twenty hours of battery life. The WTP was $113.03 when we used individual price sensitivity. When we used the average price sensitivity, the WTP dropped to $96.16. The difference would undoubtedly matter to parties in patent litigation, but the values are in the same ballpark.

[91]Because individual WTP values did not have normal bell-shaped distribution, we calculated all 95 percent confidence intervals in this article using bootstrapping taking 10,000 samples. *See generally* Bradley Efron & Robert J. Tibshirani, An Introduction to the Bootstrap 45, 113, 183 (1993).

However, the 95 percent confidence intervals are dramatically different. Confidence intervals give some indication of the certainty of an estimate. Here, the 95 percent confidence interval represents the range of values that we are 95 percent certain contains the true WTP average produced by the conjoint analysis survey. Using individual price sensitivities yields a 95 percent confidence interval for twenty hours of battery life, ranging from –$4.40 to $231.27. Because the range dips below zero, the results of the conjoint analysis cannot even show with certainty that the feature has *any* value. A jury is unlikely to give much weight to such a finding in determining damages. In contrast, using the average price sensitivity yields a 95 percent confidence interval, ranging from $96.14 to $96.19. Because the range is so tight, an expert could testify that she was highly confident in the results of the conjoint analysis. These tight confidence intervals tell juries that the results are both reliable and precise.

This pattern repeats itself for all of the features. In Figure 4, all of the 95 percent confidence intervals calculated using individual price sensitivities are disturbingly large. Moreover, the confidence intervals for three of the four kinds of features (all but noise reduction) dip below zero, suggesting that the results are not trustworthy. In contrast, the 95 percent confidence intervals calculated using average price sensitivities are extremely tight. The precise values can be found in Tables B1 and B2 of the Appendix. While Figure 4 uses only the results of the Low-value Conjoint survey to show the difference between using individual and average price sensitivity, the same differences appear when viewing the results of the Low-value Conjoint survey (Appendix, Tables C1 and C2).

C. Trimming Techniques

Experts may also choose either to trim or winsorize data, two techniques that are used to deal with outliers.[92] Trimming refers to discarding outliers from the data set. For example, an expert might eliminate the top 5 percent and bottom 5 percent of the data set. Winsorizing is similar to trimming except that instead of discarding the data, the outliers are transformed to specified upper and lower bounds.

---

[92] ALAN REIFMAN & KRISTINA KEYTON, ENCYCLOPEDIA OF RESEARCH DESIGN, 1636–37 (Neil J. Salkind ed., 2010).

Both approaches are commonly used, and there are many possible variations (e.g., treating only high outliers, using different thresholds). There may be good reasons for using these techniques. For example, a study may ask respondents to put a price on being wrongfully fired. The researcher may believe that the respondents who gave excessively high responses (e.g., over $50,000,000) did not take the tasks seriously. In that case, the researcher could discard responses over a particular threshold (i.e., trim). Alternatively, the researcher may think respondents who provided excessively high values were expressing outrage but overstating the actual price. In this case, the researcher may choose to transform any responses that exceeded some upper limit (i.e., winsorize).

While these approaches have their place, it is never appropriate to pick and choose among these techniques to obtain a desired result.[93] We demonstrate the pernicious effects of "data fishing" here by doing just that, and we find approaches that favor either the patentee or the infringer. Our data includes a number of outliers. Specifically, we examine three different approaches for handling outliers: (1) keeping all the data;[94] (2) trimming the data by dropping 5 percent of the outliers on both the top and bottom of the data set; and (3) trimming responses with positive price sensitivities. The third approach eliminates a category of unrealistic responses, namely, those responses that suggest that the respondent prefers to pay higher pricing for the same features.[95]

Figure 5 below depicts the results of applying these three approaches to the data from our low-value experiment.[96] We focus on the features tested in both experiments, battery life and location privacy. An expert testifying

---

[93]Bavli, *supra* note 70, at 11 (explaining how a researcher may improperly winsorize data simply to achieve a particular result).

[94]Again, we discarded one outlier for the low value. *See supra* note 90.

[95]*See* Orme, *supra* note 32, at 85 ("Another problem arises when price coefficients are positive rather than negative as expected. This may happen for some respondents due to random noise in the data or respondents who are price insensitive. Such reversals would suggest willingness to pay more for less desirable features."); Visteon Glob. Technologies, Inc. v. Garmin Int'l, Inc., No. 10-CV-10578, 2016 WL 5956325, at *5 (E.D. Mich. Oct. 14, 2016) ("Dr. Steckel constrained the partworths on the price attribute for each respondent to be monotonic, i.e. his analysis imposed the reasonable economic assumption that individuals prefer paying a lower price for a product than a higher price for that same product.").

[96]The underlying data from Figure 5 is found in Appendix Tables D1–D3.

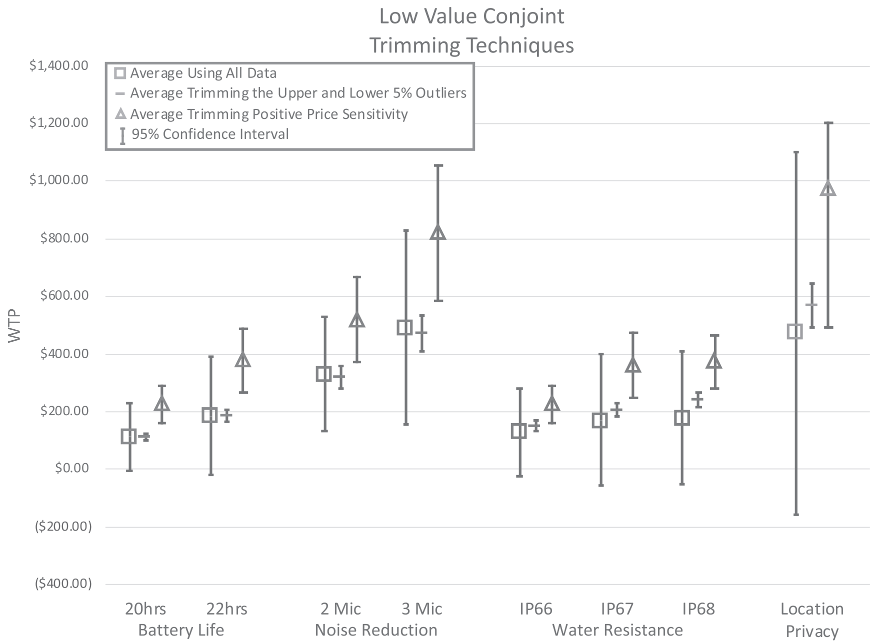


Figure 5. Different Trimming Techniques

on behalf of the patent owner could choose the approach that maximizes values. Here, that analysis clearly drops all positive price sensitivities. The value of twenty-two hours of battery life (as opposed to nineteen hours) was $379 and the value of location privacy (as opposed to the wireless carrier selling location information) was $976. On the other hand, an infringer could choose the approach that minimizes prices. Here, that approach retains all data. The resulting values are $185 for twenty-two hours of battery life and $475 for location privacy. We recognize an infringer would never offer evidence suggesting such unrealistically high prices. However, our point is that these approaches yield very different outcomes: a $194 difference in the case of battery life and a $501 difference in the case of location privacy. That our choices yielded large swings in WTP values exposes a serious danger. Experts could use all of these approaches but report only the most favorable results. Again, Figure 5 provides a comparison of trimming techniques using the results from only the Low-value Conjoint, but the results from the High-value Conjoint suffer similar problems (Appendix, Tables E1–E3).

D. Manipulating Other Features

We hypothesized that a conjoint analysis survey that included low-value features with the features of interest would lead to a higher valuation than a conjoint analysis survey that included relatively high-value features. Accordingly, we tested two common features (battery life and privacy) in a first survey with two low-value features (background noise reduction and water resistance) and in a second survey with two high-value features (screen resolution and storage).

We calculated the difference between the average WTP for the features of interest using bootstrapping. Table 5 below depicts both the average difference of means and the basic bootstrap 95 percent confidence interval for the difference of the means. We hypothesized that the lower value survey would increase the value of these features. The results for the battery life feature were consistent with the hypothesis. On average the WTPs for twenty and twenty-two hours of battery in the low-value experiment were worth \$19.38 and \$50.91 more than in the high-value experiment respectively. Because the 95 percent bootstrap confidence intervals did not fall below zero, the results show bootstrapped statistical significance at the $p < 0.05$ level.

However, our results for location privacy were less conclusive. On average the WTP for location privacy in the low-value experiment was worth \$11.15 more than in the high-value experiment. However, the 95 percent confidence interval fell between negative \$83.89 and positive \$108.06. Accordingly, there is little evidence of replicability of the greater WTP result for this feature in the low-value experiment.

At first blush, our findings are not fully consistent with Bedi and Reibstein's more dramatic results. They found that the value of more cup holders was significantly higher when the conjoint analysis survey included lower value features as opposed to higher value features.[97] There are several potential ways to interpret the studies together. First, our study focused on two relatively important features, battery life and location privacy. In contrast, Bedi and Reibstein tested a fairly trivial set of features, increasing the number of cup holders in a car and the location of a car's gas cap. It may well be that using less valuable features in a conjoint analysis only inflates prices when the features of interest are not

[97] *See supra* notes 56–60 and accompanying text.

valuable either. Second, even though our three results had varying degrees of statistical significance (twenty and twenty-two hours of battery life significant and location privacy insignificant), the results all pointed in the same direction as Bedi and Reibstein's results. Thus, both studies may tend to prove that including low-value features inflates the value of the feature of interest. However, the degree of inflation may somehow be related to the feature of interest's relative importance to the other features used in the survey. Finally, the results of the two studies may be simply inconsistent. Although our intuition is that some version of the first or second interpretation is correct, further research would be needed to resolve this question.

V. Policy Recommendations

A. Monetary Values

First, our results reinforce recommendations made by others.[98] At least in patent lawsuits involving numerous features, conjoint analysis should not be used to place a dollar value on the infringing feature. Conjoint analysis dramatically inflates feature values in ways that are impossible to justify. Given the paucity of alternatives, we recognize that courts might be tempted to allow experts to introduce data from a conjoint analysis survey and trust juries to identify and assess the flaws in this approach. We maintain that courts should resist this temptation. Like all people, jury members are subject to heuristics like anchoring. In study after study, researchers have shown that an initial damage demand, no matter how baseless, can influence the jury's decision.[99] Thus, the introduction of specific WTP valuations based on conjoint analysis can significantly

---

[98] *See* Orme, *supra* note 32, at 85 (discouraging the use of conjoint analysis to monetize features because the analysis "is a potentially misleading."); Bedi & Reibstein, *supra* note 32, at 50 (noting problems when conjoint analysis is used to determine relative importance "in terms of dollars").

[99] Chao & O'Dorisio, *supra* note 46, at 35 ("For now, we still have not identified a particularly effective response to a plaintiff that anchors the jury in an irrationally high number."); Gretchen B. Chapman & Brian H. Bornstein, *The More You Ask for, the More You Get: Anchoring in Personal Injury Verdicts*, 10 Applied Cognitive Psychol. 519, 522 (1996). *See generally* John Campbell, Bernard Chao, Christopher Robertson & David V. Yokum, *Countering the Plaintiffs' Anchor: Jury Simulations to Evaluate Damages Arguments*, 101 Iowa L. Rev. 543 (2016).

impact the jury's decision even if the jury is convinced that the underlying approach is deeply flawed. Worse yet, unlike other types of crude estimates that are often used, conjoint analysis has the veneer of scientific precision. Thus, there is a danger that juries might give more weight to conjoint analysis than they would in calculating die area, feature counts, or line counts.

B. Relative Valuations

A more reliable approach would be to allow experts to use conjoint analysis to suggest the relative importance of features. The court in *Visteon Glob. Technologies v. Garmin Int'l* explained how such an approach might work.[100] A conjoint analysis could assess the value of both patented and other important nonpatented features.[101] This would provide some relative sense of worth. In other words, even if conjoint analysis does not yield accurate WTP values, it may help to determine whether a patented feature is worth "substantially more" or "substantially less" than other important features. At this level of generality, we have no objection to the use of conjoint analysis to help juries get a sense of the value of a patented feature.

However, experts may be tempted to give conclusions with a misleading level of precision. For example, one might interpret our results to suggest that location privacy is 10.7 times as valuable as a 128 GB increment of storage.[102] Because storage has established market values, we might conclude that we simply need to multiply the value of 128 GB by 10.7 in order to arrive at the value of location privacy.

Our experiments do not support this approach. While conjoint analysis yielded absurdly high results, the magnitude of the problem was not uniform. We could not say that the values for each feature were off by a single specific multiple. In fact, for one feature, storage, the results appeared too low. This dynamic makes it impossible to solve the problem by scaling the conjoint analysis results downward by a given multiple. Moreover, this is simply another way to try to obtain a level of precision

---

[100] Visteon Glob. Technologies, Inc. v. Garmin Int'l, Inc., No. 10-CV-10578, 2016 WL 5956325, at *6 note 5 (E.D. Mich. Oct. 14, 2016).

[101] *Id*.

[102] This calculation is derived by taking $558 (the value of location privacy) and dividing it by $52 (the value of 256 GB as compared with 128 GB). *See* Appendix Table A.

that conjoint analysis does not provide. This is essentially the same problem as trying to place a monetary value on a tested feature.

However, we recognize the suggestion that conjoint analysis gives some reliable sense of relative valuations.[103] Therefore, there is support for permitting experts to use conjoint analysis in order to aid juries in providing some rough relative valuations. Even then, we urge caution. Complex features should not be valued through surveys that treat the feature too superficially. For example, others have shown that privacy valuations typically have huge fluctuations due, in part, to the respondents' failure to understand what privacy is.[104] Many patented features are also extremely opaque, making it unlikely that survey respondents will be able to grasp the subtleties and implications of the technology in the short time they have to answer a series of survey questions. If the respondents do not adequately comprehend the feature (as compared with its non-infringing alternative), we can expect that valuations based on their answers will be ill-informed, and therefore unreliable.

C. Confidence Intervals/Price Sensitivity

If a court does allow an expert to introduce the results of a conjoint analysis experiment, we recommend that it require the expert to include confidence intervals when providing her results. For social scientists, this recommendation may appear overly rudimentary. However, we found that while some experts express their results using confidence intervals, others simply characterize their results as statistically significant, typically by using *p*-values or saying nothing at all.[105]

Few if any studies provide 100 percent proof of a proposition. Accordingly, *p*-values are helpful for understanding how well the results of a study prove or fail to prove a given hypothesis (e.g., a vaccine reduces the

---

[103]Orme, *supra* note 32, at 26 (discussing conjoint analysis as providing "directional indicators"); Bedi & Reibstein, *supra* note 32, at 50 (suggesting that conjoint analysis can estimate the relative preference of features).

[104]*See supra* note 63.

[105]*See, e.g.*, Plaintiff's Exhibit No. 30, Apple, Inc. v. Samsung Elecs. Co., No. 11-CV-01846-LHK, 2012 WL 6115623 (N.D. Cal. Aug. 2, 2012) (demonstrative exhibit providing WTP values without confidence interval or *p*-values). Professor Hauser's trial testimony also did not contain this information. Trial Transcript pp. 1914–17 (Aug. 10, 2012).

incidence of later acquiring a disease).[106] The value corresponds to the probability that the same results could have been arrived at by chance. Thus, a low *p*-value corresponds to strong evidence, and most disciplines consider *p*-values of less than .05 (i.e., less than a one in twenty chance) to be statistically significant.[107] However, *p*-values do not show the magnitude of an effect (e.g., how much does the vaccine reduce the likelihood of acquiring a disease?). It is here that confidence intervals can help. A confidence interval suggests the range of possibilities that are consistent with the results. Typically, 95 percent confidence intervals are used because that range corresponds to a *p*-value of .05, statistical significance.

We recommend that when experts provide the specific results of any conjoint analysis survey, they also provide a confidence interval. For example, a conjoint analysis survey may reveal that consumers value feature A at $36. If the expert provides only that number, the jury may mistakenly think that $36 is "the" value of feature A. But if the expert also tells the jury that the 95 percent confidence interval is between $36 and $46, juries are more likely to understand that the results of the conjoint analysis survey are consistent with a range of possible values. Moreover, as the range gets larger (e.g., $16 to $56), the jury might appropriately start questioning how reliable the results are. Confidence intervals are regularly used to describe the results of scientific studies, and experts should have no problem providing such information. While some may be concerned that jurors will not understand confidence intervals, studies have shown that juries can correctly incorporate straightforward statistical evidence into their decisions.[108] Accordingly, this recommendation should be uncontroversial.

---

[106] *See* David H. Kaye & David A. Freedman, Reference Manual on Scientific Evidence, Reference Guide on Statistics 252–53 (3rd ed. 2011) ("By inquiring into the magnitude of an effect, courts can avoid being misled by *p*-values. To focus attention on more substantive concerns—the size of the effect and the precision of the statistical analysis—interval estimates (*e.g.*, confidence intervals) may be more valuable than tests."); *see generally* Jean-Baptist du Prel, Gerhard Hommel, Bernd Rohrig & Maria Blettner, *Confidence Interval or p-Value?*, 106 Deutsches Arzteblatt Int'l 335 (2009).

[107] Kaye & Freedman, *supra* note 106, at 251.

[108] *See generally* Rebecca K. Helm, Valerie P. Hans & Valerie F. Reyna, *Trial by Numbers*, 27 Cornell J.L. & Pub. Pol'y 107 (2017) (reviewing numerous studies on jurors' ability to understand statistical evidence and suggesting potential ways to help jurors address potential sources of error); Brian C. Smith, Steven D. Penrod, Amy L. Otto & Roger C. Park, *Jurors' Use of Probabilistic Evidence*, 20 Law & Hum. Behav. 49, 78 (1996) (finding that mock jurors generally made reasonable use of probabilistic evidence, but with great variation).

D. Pre-File Methodologies

Like many types of scientific analyses, conjoint analysis is susceptible to manipulation.[109] As our examples illustrate, experts can perform many different types of otherwise reasonable calculations and conveniently select the one that best suits their interests.[110] To avoid this problem, we recommend that judges insist that damages experts file a detailed description of their methodological approach with the court before they run any conjoint analysis survey and analyze the data. For experts who plan to use conjoint analysis, the filing should describe all of the study's important attributes, including (1) what features will be included in the survey, (2) the different levels of the feature, (3) how they will be described, (4) how participants will qualify or disqualify, and (5) how outliers will be handled. Ideally, the filing should not be sealed.[111] This transparency will allow the analysis to be subject to peer review, which could lead to better analysis over time.

The idea of preregistering a study is already considered a "best practice" in the scientific community. For example, the *New England Journal of Medicine* publishes statistical guidelines for researchers who plan to submit an article. The guidelines recommend that for articles that report on clinical trials, "final protocols and statistical analysis plans (SAPs) … be submitted along with manuscript."[112] Hillel Bavli has recommended that courtroom experts use similar protocols.[113] However, Bavli does not

---

[109]*See* Bavli, *supra* note 70, at 20–28; Jonah B. Gelbach, *Expert Mining and Required Disclosure*, 81 U. Chi. L. Rev. 131, 135–36 (2014) (explaining how experts can conduct numerous tests but select the one that produces the most favorable results).

[110]Of course, this description applies to many kinds of expert analyses and our recommendation makes sense beyond the damages world. But for now, we focus on conjoint analysis.

[111]*See generally* Bernard Chao & Derigan Silver, *A Case Study in Patent Litigation Transparency*, 2014 J. Disp. Resol. 83 (2014) (discussing the problem of oversealing filings in patent cases).

[112]*New Manuscripts*, New Eng. J. Med, https://www.nejm.org/author-center/new-manuscripts (last visited Dec. 28, 2020); *see also* Kai Kupferschmidt, *More and More Scientists Are Preregistering Their Studies. Should You?*, Science, Sept. 21, 2018, https://www.sciencemag.org/news/2018/09/more-and-more-scientists-are-preregistering-their-studies-should-you (stating that preregistration has become the "norm for clinical trials").

[113]Bavli, *supra* note70, at 33 ("[T]he researcher should specify her design in writing and record it, ideally, by publishing it, "preregistering" it, or at least privately uploading it electronically with a timestamp.").

advocate for courts to require preregistration.[114] Instead, he merely suggests that courts consider whether the protocols should be preregistered in assessing an expert's credibility.[115] In a similar vein, Jonah Gelbach believes that reliance on a combination of expert discovery and *Daubert* rules can adequately deal with experts who engage in data fishing.[116]

We are less optimistic than Gelbach. Experts, particularly in patent cases, are often professional witnesses. They are able to deliberately obfuscate and delay. It is unclear whether even well-prepared attorneys will be able to discover all but the most obvious cases of data fishing. Moreover, judges and juries may have a hard time understanding why an otherwise reasonable approach should be ignored simply because an expert chose that approach after looking at the results of all of her possible choices.

Accordingly, we believe that stronger medicine is called for here, and that is why we recommend that courts make preregistration mandatory. For litigation, it should mean filing a research plan with the court. The courts' inherent power to manage litigation already gives them authority to create local rules or issue standing orders to implement this recommendation.[117] If a court does not have such a rule, attorneys should consider asking for preregistration in their case management conference.[118] Once there is a written record of the expert's plan, it will be far easier for opposing counsel to learn whether an expert has strayed from that plan. If an expert does not preregister her work, she should not be able to introduce findings later. To the extent that the expert pre-files but

---

[114]Bavli, *supra* note 70, at 43 ("My intention is not to argue that courts should necessarily exclude all analyses for which the researcher is unable to establish strict adherence to the principles … .").

[115]Bavli, *supra* note 70, at 43.

[116]Gelbach, *supra*, note 109, at 136. ("[T]he Federal Rules of Civil Procedure and of Evidence work together to enable a well-prepared party to punish an adversary for its expert's use of data mining.").

[117]Fed. R. Civ. P. 83(a) authorizes district courts to adopt local rules while Rule 83(b) authorizes judges to create rules of practice for their courts. *See also* Justin Sevier, *Unintended Consequences of Local Rules*, 21 Cornell J.L. & Pub. Pol'y 291, 301–02 (2011) (discussing the basis for procedural local rules).

[118]Fed. R. Civ. P. sixteen provides for a pretrial conference to create a "scheduling order" that may modify the timing and extent of discovery.

changes her approach, she should have to justify the changes or risk exclusion.

Finally, if the other party believes that the approach is unfair, it can identify problems and suggest revisions. This approach may have the added benefit of introducing studies that are not merely designed to benefit one side or the other, but instead, take a more balanced view.[119] Allowing experts to revise their study's methodology before implementation also gives courts better options. Currently, when courts see flaws with an expert's analysis, they have limited choices. They can (1) exclude the report leaving the party with no expert testimony, (2) give the expert a do-over, or (3) allow the testimony, despite its reservations, with the belief that the jury will see the flaws, and if not, grant some kind of post-trial motion.[120] As the Patent Case Management Judicial Guide notes, these choices are all "imperfect options."[121] The court either leaves a party entirely without an expert, wastes resources, or gives the jury expert evidence with known flaws. In contrast, requiring experts to fix problems earlier saves resources and provides the jury with more reliable evidence that is less likely to present extreme conclusions.

Of course, this recommendation is equally applicable to many types of analyses that testifying experts regularly perform inside and outside of patent law. Accordingly, we believe that courts should consider requiring experts to pre-file their approaches in other contexts as well.

Conclusion

In an effort to determine whether conjoint analysis is a useful tool to evaluate the value of specific features in complex multicomponent patent lawsuits, this article discusses the results of two conjoint analysis surveys. For the most part, the surveys yielded unrealistically high values. Accordingly, we recommend that courts do not allow experts to offer the results of conjoint analysis as evidence of specific monetary values. However,

---

[119] *See* Christopher Tarver Robertson, *Blind Expertise*, 85 N.Y.U. L. Rev. 174, 189 (2010) (noting that currently juries see "only two opposite extremes of expert opinion" instead of opinions that lie in between which may be closer to the truth.).

[120] Peter Menell et al., Patent Case Management Judicial Guide, § 2.6.6 (3rd ed. 2016) (citations omitted).

[121] *Id*.

there is some reason to believe that conjoint analysis can be used to provide relative valuations (i.e., feature A is worth substantially more than feature B). Even then, we urge caution and explain how data from conjoint analysis surveys can be manipulated and why the approach might not be appropriate for some hard-to-understand features. To the extent that courts allow the use of conjoint analysis, we provide several recommendations to prevent manipulation.

Appendix[122]

Table A: Valuations Using Average Price Sensitivities and 10 Percent Trim

	High Value			*Low Value*		
Features	*Mean*	*Lower Bound CI*	*Upper Bound CI*	*Mean*	*Lower Bound CI*	*Upper Bound CI*
Battery (20 hours)	$92.35	$85.24	$99.50	$111.73	$98.71	$124.68
Battery (22 hours)	$135.63	$124.33	$146.90	$186.54	$164.21	$208.61
Location Privacy	$557.97	$498.85	$617.02	$569.11	$493.82	$645.59
Screen Res (1080p)	$295.90	$266.54	$324.97			
Screen Res (2160p)	$603.97	$542.87	$665.69			
256 GB	$51.61	$45.36	$57.78			
512 GB	$132.40	$122.20	$142.71			
1T	$220.66	$202.82	$238.24			
Background Noise (2 Mic)				$320.91	$280.03	$361.33
Background Noise (3 Mic)				$470.68	$407.93	$534.47
Water Resistance (IP66)				$151.35	$132.02	$170.39
Water Resistance (IP67)				$206.54	$183.09	$230.42
Water Resistance (IP68)				$241.86	$216.00	$267.55

[122]All 95 percent confidence intervals in this appendix were calculating by bootstrapping (10,000 samples).

Table B1: Low Value, Willingness to Pay Results Using Average Price Sensitivity

Individual Price Sensitivities	*Battery (20 hours)*	*Battery (22 hours)*	*Background Noise (2 Mic)*	*Background Noise (3 Mic)*	*Water Resistance (IP66)*	*Water Resistance (IP67)*	*Water Resistance (IP68)*	*Data Privacy*
Mean	$113.03	$185.34	$329.28	$489.36	$131.75	$169.57	$177.67	$474.89
Lower CI	($4.40)	($17.65)	$131.40	$154.42	($21.79)	($56.91)	($53.01)	($156.48)
Upper CI	$231.27	$389.84	$527.41	$827.13	$281.27	$401.20	$409.74	$1102.58

The top 5 percent and bottom 5 percent of the results based on price sensitivity were trimmed on in calculating the results in Table B1, B2, C1, C2.

Table B2: Low Value, Willingness to Pay Results Using Individual Price Sensitivity

Average Price Sensitivities	*Battery (20 hours)*	*Battery (22 hours)*	*Background Noise (2 Mic)*	*Background Noise (3 Mic)*	*Water Resistance (IP66)*	*Water Resistance (IP67)*	*Water Resistance (IP68)*	*Data Privacy*
Mean	$96.16	$160.77	$275.15	$399.35	$135.01	$186.10	$215.08	$501.05
Upper CI	$96.19	$160.85	$275.35	$399.88	$135.11	$186.26	$215.29	$501.90
Lower CI	$96.14	$160.69	$274.94	$398.82	$134.90	$185.94	$214.86	$500.19

Table C1: High Value, Willingness to Pay Results Using Average Price Sensitivity

Individual Price Sensitivities	*Battery (20 hours)*	*Battery (22 hours)*	*Screen Res (1080p)*	*Screen Res (2160p)*	*256 GB*	*512 GB*	*1T*	*Data Privacy*
Mean	$111.45	$171.19	$294.37	$667.98	$63.28	$149.14	$260.84	$355.40
Lower CI	$40.46	$58.62	($58.62)	($84.25)	$43.79	$60.77	$105.85	($687.65)
Upper CI	$183.46	$284.08	$643.77	$1404.04	$82.47	$237.60	$417.89	$1393.62

Table C2: High Value, Willingness to Pay Results Using Individual Price Sensitivity

Average Price Sensitivities	*Battery (20 hours)*	*Battery (22 hours)*	*Screen Res (1080p)*	*Screen Res (2160p)*	*256 GB*	*512 GB*	*1T*	*Data Privacy*
Mean	$83.64	$122.89	$251.64	$506.99	$45.70	$115.97	$191.56	$507.84
Upper CI	$84.21	$125.67	$253.46	$518.09	$46.07	$117.27	$194.52	$525.74
Lower CI	$83.06	$120.12	$249.82	$495.89	$45.33	$114.66	$188.59	$489.95

Table D1: Low Value, Willingness to Pay Results Using All Data

All Data	*Battery (20 hours)*	*Battery (22 hours)*	*Background Noise (2 Mic)*	*Background Noise (3 Mic)*	*Water Resistance (IP66)*	*Water Resistance (IP67)*	*Water Resistance (IP68)*	*Data Privacy*
Mean	$113.03	$185.34	$329.28	$489.36	$131.75	$169.57	$177.67	$474.89
Lower CI	($4.40)	($17.65)	$131.40	$154.42	($21.79)	($56.91)	($53.01)	($156.48)
Upper CI	$231.27	$389.84	$527.41	$827.13	$281.27	$401.20	$409.74	$1102.58

Individual Price Sensitivity was used to calculate the results in Tables D1–D3. In addition, one respondent with an individual price sensitivity three orders of magnitude above the median and exclusive mean was excluded from results for Tables C1, C2, and E1–E3 (all). *See supra* note 90.

Table D2: Low Value, Willingness to Pay Results 10 Percent Trim (top and bottom)

Winnow + 5% & -5%	*Battery (20 hours)*	*Battery (22 hours)*	*Background Noise (2 Mic)*	*Background Noise (3 Mic)*	*Water Resistance (IP66)*	*Water Resistance (IP67)*	*Water Resistance (IP68)*	*Data Privacy*
Mean	111.728	186.537	320.906	470.677	151.352	206.537	241.860	569.112
Lower CI	98.705	164.211	280.034	407.928	132.020	183.094	215.995	493.821
Upper CI	124.676	208.614	361.327	534.465	170.388	230.416	267.551	645.588

Table D3: Low Value, Willingness to Pay Results Using Only Positive Price Sensitivities

Positive Data	*Battery (20 hours)*	*Battery (22 hours)*	*Background Noise (2 Mic)*	*Background Noise (3 Mic)*	*Water Resistance (IP66)*	*Water Resistance (IP67)*	*Water Resistance (IP68)*	*Data Privacy*
Mean	225.744	379.078	519.874	820.426	225.744	360.631	373.160	975.875
Lower CI	159.967	268.255	373.878	587.097	159.967	248.317	280.715	493.821
Upper CI	290.847	490.206	666.887	1055.732	290.847	472.324	465.246	1203.255

Table E1: High Value, Willingness to Pay Results Using All Data

All Data	*Battery (20 hours)*	*Battery (22 hours)*	*Screen Res (1080p)*	*Screen Res (2160p)*	*256 GB*	*512 GB*	*1T*	*Data Privacy*
Mean	$111.45	$171.19	$294.37	$667.98	$63.28	$149.14	$260.84	$355.40
Lower CI	$40.46	$58.62	($58.62)	($84.25)	$43.79	$60.77	$105.85	($687.65)
Upper CI	$183.46	$284.08	$643.77	$1404.04	$82.47	$237.60	$417.89	$1393.62

Individual Price Sensitivity was used to calculate the results in Tables E1–E3.

Table E2: High Value, Willingness to Pay Results 10 Percent Trim (top and bottom)

Winnow + 5% & -5%	*Battery (20 hours)*	*Battery (22 hours)*	*Screen Res (1080p)*	*Screen Res (2160p)*	*256 GB*	*512 GB*	*1T*	*Data Privacy*
Mean	$92.35	$135.63	$295.90	$603.97	$51.61	$132.40	$220.66	$557.97
Lower CI	$85.24	$124.33	$266.54	$542.87	$45.36	$122.20	$202.82	$498.85
Upper CI	$99.50	$146.90	$324.97	$665.69	$57.78	$142.71	$238.24	$617.02

Table E3: High Value, Willingness to Pay Results Using Only Positive Price Sensitivities

Positive Data	*Battery (20 hours)*	*Battery (22 hours)*	*Screen Res (1080p)*	*Screen Res (2160p)*	*256 GB*	*512 GB*	*1T*	*Data Privacy*
Mean	$153.05	$227.56	$528.97	$1161.01	$60.82	$210.14	$372.60	$973.47
Lower CI	$93.85	$127.28	$335.49	$709.94	$46.42	$148.44	$259.04	$516.44
Upper CI	$212.45	$328.82	$721.68	$1609.36	$75.21	$272.23	$486.56	$1427.23