Product Owner Surveys

Consumer Reports' product ratings are unique in that they combine product performance, measured through expert lab testing, with predicted reliability and owner satisfaction data gathered through surveys of product owners. The latter two measures are based on consumers' real-world experiences with products they have purchased and reflect consumers' use of products over time. By combining lab performance testing with data based on the experiences of hundreds of thousands of real consumers, Consumer Reports ratings give consumers the most complete information on what they can expect from each product.

Data on product reliability and owner satisfaction is gathered through regular surveys of Consumer Reports members. In these surveys, respondents answer questions about the products they own—in that way, Consumer Reports members are true partners in our product rating process! Survey questions typically include how long the respondent has owned a particular product, how often they use it, whether it has broken or stopped working as well as it should, what types of problems the product has had, and whether they would recommend the product to someone else. We also ask survey participants to tell us in their own words what they do and do not like about the product they own and to describe in more detail any problems they may have experienced.

In a typical year we invite roughly 3 million Consumer Reports members to participate in our online product surveys, and gather product reliability and satisfaction data in up to 40 product categories. Our expert team of survey methodologists and data analysts use that data to calculate the predicted reliability and owner satisfaction scores you see for each brand in a specific product category. Our surveys are a powerful and unique tool, enabling us to gather data from a large enough sample of consumers to be able to rate specific brands across a large number of product categories.

As with all the research and testing Consumer Reports does, our surveys are designed to gather unbiased, objective information from consumers and our analyses are conducted free of any corporate or industry influence. Read here for more about how our member surveys are conducted and how the resulting data is used to calculate product reliability and owner satisfaction for all product categories other than autos. A description of how reliability is calculated for autos can be found here.

Fielding CR Member Surveys

- Consumer Reports' reliability ratings are based on the real-world experiences of consumers with the products they own. To calculate predicted reliability for products, we

regularly survey CR members to find out whether certain products they own have broken or had problems. We use the resulting data on problem rates to estimate how likely new models from a given brand in that product category are to hold up over time.

- Our member surveys are fielded online each quarter, with each quarter focusing on a different set of up to 10 product categories. Each quarterly survey is accessed via an email invitation sent to eligible CR members. To be eligible to receive a survey invitation, CR members must have a valid email address in our files. If you are a member, you can update your email address here.

- Not all eligible CR members are invited to take each quarterly survey. While we do invite all eligible members to complete our Spring Survey each year because it includes our Annual Auto Survey, our Winter, Summer, and Fall Surveys are sent to a random sample of eligible members.

- Member surveys remain in the field for about seven weeks each quarter, and members receive multiple reminders to complete survey sections for the products they own. Each product section is treated as a separate survey, meaning members do not have to complete all sections for their data to be used.

- CR members do not receive incentives for completing surveys or survey sections.

Calculating Predicted Reliability

- Depending on the product category (and what makes the most sense for each product we rate), our predicted reliability estimates are based on responses to one of the following three survey questions:

  - *Did this product ever break or stop working as well as it should?* (Applicable for all major appliances, many electronic devices, and outdoor power tools.)

  - *Have you ever had problems with this product since you've owned it?* (Applicable for small appliances, including blenders and coffee makers.)

  - *Which of the following best describes the reliability of this product during the time you have owned it?* (Applicable for printers and smartphones; response choices are "very reliable," "somewhat reliable," "somewhat problematic," and "very problematic.")

- Predicted reliability ratings are based on estimated instances of problems occurring within a given time frame, or "problem rates." Problem rates within each product category are calculated at the brand level, not for each specific product model. So, for example, when rating dishwashers, we calculate problem rates for each brand as a whole (Bosch, GE, KitchenAid, Whirlpool, etc.) as opposed to calculating problem rates for each GE or Whirlpool dishwasher model separately. We do this for two reasons: First, while consumers can typically easily identify the brand of each product they own,

2

they have much more difficulty accurately reporting the specific model they own; second, calculating problem rates at the brand level enables us to use larger sample sizes in our calculations, thereby producing more precise estimates for a larger range of products.

- For product categories with multiple major subcategories (for example, various types of refrigerators, such as French-door, top-freezer, bottom-freezer, side-by-side, and built-in), problem rates are calculated at the brand level but are subcategory-specific. This means, for example, that the reliability rating we ultimately give to Brand X's top-freezer refrigerators is calculated separately from the rating we give to Brand X's bottom-freezer refrigerators.

- CR member surveys do not just ask consumers about purchases made in that survey year. Rather, because our goal is to measure product experience over time, each survey gathers data on products purchased across a longer time frame. The number of years of purchases for which we gather data and then include in our reliability analyses varies by product category and is based on members' expectations of how long newly purchased products in that category should last (as gathered through separate survey questions). Currently:

    - Ten years is members' average expected life span for all major appliance categories.

    - Eight years is members' average expected life span for blenders and coffee makers.

    - Five years is members' average expected life span for computers.

    - Three years is members' average expected life span for smartphones.

- Advanced statistical models are used to calculate a brand's predicted reliability rating (a reflection of its estimated problem rate) in a particular product category as a function of product age (measured through year of purchase), frequency of use, and extended warranty or service contract coverage. We control for these factors in our models to ensure objective comparisons among brands.

- Across all product categories, predicted reliability ratings are set at the midlife of a product's expected life span. So for major appliances, our reliability predictions are based on the likelihood of problems occurring within the first five years because our members expect to get 10 useful years of ownership out of such products, on average. We predict to the middle year rather than the last year because estimates that are pinpointed to the middle year are more reliable statistically than those made to the final year of the expected life span, and, because most products will have experienced one or more problems by the final years of the life span, there is less reliability differentiation across brands at that point. In addition, our analyses consistently show that if a brand

appears to be problematic in a product category by the middle year of the life span, it is a good indication that the brand will be even more problematic in later years.

- ○ We should note that in our surveys, we ask our members to report on <u>all purchases</u> made within a specified time frame, regardless of whether they still own those products. Depending on the average expected life span of a particular product category, the time frame we ask about can be purchases made within the past three years, past five years, past eight years, or past 10 years. We include <u>previously owned products</u> in addition to those currently owned, which ensures that our reliability prediction models account for products that failed quickly and/or did not meet average expectations of product longevity.

- If we have sufficient data for a brand, we assign it its own reliability rating for that product category. If we do not have adequate data to give a brand its own rating, we take the following steps:

  - ○ First, we look to see whether we have enough data on other brands in that product category belonging to the <u>same parent company</u>.

  - ○ If we do, we base our prediction for the brand lacking in adequate sample on an average of how the other brands belonging to the same parent company are rated in that category.

  - ○ Barring that, we assign brands lacking adequate sample a <u>category average rating</u>. When this occurs, we display a dash in CR ratings tables to indicate that we did not have adequate data to calculate a reliability prediction for that brand.

  - ○ In these cases, while the predicted reliability rating is the category average, if we do have any sample from that brand, the data is weighted as part of the predicted reliability score factored into the Overall Score.

Assigning Brand Reliability Ratings

- Predicted reliability, like all of CR's rated product attributes, is classified into five ratings tiers, where 1 is the worst and 5 is the best. The problem-rate thresholds that delineate one ratings tier from another are product category-specific and are based on data collected from our most recent surveys fielded over the past two or three years. With this approach we can account for market shifts that may occur from year to year in technology, innovation, and consumer expectations within each product category as we assign reliability ratings.

- Thresholds between reliability tiers are set in such a way as to identify brands that are <u>more or less reliable than most of their peers</u> on the market at that time. We calculate ratings tier thresholds using standardized Z-scores and distance from the category's reliability score mean in a given year. This enables us to identify brands that are

performing much better or worse than their peers in each given year and enables manufacturers to determine whether they are continuing to improve or deteriorate over time relative to competitors. Brands that get a 4 (light green) or 5 (dark green) rating for reliability are more reliable than the average problem rate for a product category, and brands that get a 1 (red) or 2 (orange) rating for reliability are less reliable than the category's average rate.

- For a brand to change its ratings tier from its previous year's predicted reliability rating (moving either up or down the five-tier scale), the new rating must be sufficiently different from the prior rating to ensure reasonable confidence that the movement is due to an actual change in reliability and not a result of statistical fluctuations. For this reason, we establish "buffer zones" around each threshold beyond which the new rating must extend to change a brand's reliability tier.

Calculating Owner Satisfaction

- Our owner satisfaction ratings are based on how likely CR members are to recommend the products they own to their friends and family, and reflect the proportion of members who are extremely likely to do so.

- The survey question that underpins our owner satisfaction ratings is worded as follows: *How likely is it that you would recommend your [product] to your friends or family?* Responses range from "extremely likely" to "not at all likely."

- This behavioral measure of owner satisfaction reflects a more actionable account of consumer sentiment regarding the products people own than asking how happy are unhappy they are with a product.

- CR's brand-level owner satisfaction ratings are based on the same purchase years as the predicted reliability ratings for each product category noted above.

Predicted Reliability and Owner Satisfaction Impact on Overall Product Scores

- Typically, CR's Overall Scores for products comprise three components: predicted reliability, owner satisfaction, and test performance.

- The weight (i.e., importance) assigned to a brand's reliability rating in their products' Overall Scores varies by product category, depending on the importance that consumers place on reliability in their purchase decisions relative to performance and price (as gathered through separate survey questions).

  - Brands that receive a predicted reliability rating of 1 or 2 in a product category cannot receive CR recommendations for their products in that category, regardless of how their models perform in lab tests. This is based on CR's belief

5

that products that perform well in lab testing but do not hold up at least fairly well over time in real-world use should not be recommended to consumers.

- The weight assigned to owner satisfaction in CR's Overall Scores is the same across all product categories. Unlike predicted reliability ratings, brands that get a low owner satisfaction rating are still eligible to be recommended by CR because a consumer's happiness with the products they own is a more subjective assessment of each product than reliability and can fluctuate more over time.