User-Generated Versus Designer-Generated Products: A Performance Assessment at Muji

Working Paper --- forthcoming at the *International Journal of Research in Marketing*

Hidehiko Nishikawa, Martin Schreier, and Susumu Ogawa

Hidehiko Nishikawa
Hosei University
Faculty of Business Administration
2-17-1 Fujimi, Chiyoda-ku, Tokyo
102-8160 JAPAN
Tel: +81-3-3264-9643
Fax: +81-3-3264-9643
Email: hidehiko@hosei.ac.jp

Martin Schreier
WU Vienna University of Economics and Business
Marketing Department
Augasse 2-6, Vienna
1090 AUSTRIA
Tel: +43-01-31336-4682
Fax: +43-01-31336-732
Email: martin.schreier@wu.ac.at

Susumu Ogawa
Kobe University
School of Business Administration
2-1 Rokkodai-cho Nada-ku
Kobe Hyogo
657-8501 JAPAN
Tel: +81-78-8-3-6932
Fax: +81-78-803-6977
Email: ogawa@kobe-u.ac.jp

Acknowledgements: The authors thank Muji, in particular Satoshi Shimizu, Tsunemi Kawana, Tetsu Takahashi, Tomisaburou Hagiwara and Yukie Saito, for their support in realizing this project. They further thank Pradeep K. Chintagunta and Berend Wierenga for their feedback on earlier drafts of the manuscript as well as Marnik G. Dekimpe, the AE, and the reviewers for their helpful suggestions and guidance during the review process. Financial support is greatly acknowledged from Japan Society for the Promotion of Science KAKENHI (research grant 22330130, 24530535) and the Danish Strategic Research Council (research grant 09-067157).

User-Generated Versus Designer-Generated Products:

A Performance Assessment at Muji

Abstract

In recent years, more and more consumer goods firms have started to tap into the creative potential of their user communities to fuel their new product development pipelines. Although many have hailed this paradigm shift as a highly promising development for firms, hardly any research has systematically compared the actual market performance of user-generated products with designer-generated products. We fill this void by presenting a unique data set gathered from the Japanese consumer goods firm Muji, which has drawn on both sources of ideas in parallel in recent years. We find that user-generated products, which are generally more novel, outperformed their designer-generated counterparts on key market performance metrics. Specifically, in the first year after introduction, sales revenues from user-generated products were three times higher and gross margins were four times greater than those of designer-generated products. These effects also increased over time: after three years, the aggregate sales revenues of user-generated products were, on average, 1.25 billion yen (approximately 16 million dollars) higher, or five times greater, than the sales of designer-generated products. The corresponding average margin was an impressive 619 million yen (approximately 8 million dollars) higher, or six times greater, than the margin for designer-generated products. Finally, user-generated products were more likely to survive the three-year observation period than designer-generated products (i.e., were still on the market three years after introduction). These findings clearly favor the paradigm shift identified in marketing research and appeal to managers considering the integration of user ideas into the process of new product development. We discuss our study's limitations and identify important avenues for future research.

Keywords: user design, user-generated products, user innovation, co-creation, customer integration, crowdsourcing, market performance, new product development, idea generation

1. Introduction

For decades (if not for much longer), professional marketers, designers, and engineers rather than consumers or users have been the dominant agents in the development of new consumer products. Although it has always been imperative to listen to customers to discern emerging market trends and unmet consumer needs, design professionals employed by firms (or their subcontractors) are usually responsible for translating the resulting opportunities into new product ideas and market offerings (Cooper 2001; Crawford and Di Benedetto 2000; Ulrich and Eppinger 1995; Urban and Hauser 1993). Until very recently, firms did not consider user involvement in the process of generating ideas for new product development (NPD) to be particularly promising. For example, as Schulze and Hoegl (2008, p. 1744) note, "relying on the method of asking buyers to describe potential future products, big leaps to novel product ideas are generally not likely." One of the key assumptions underlying this dominant idea generation paradigm is that a firm's professionals, unlike customers or users, have the requisite expertise to invent new and useful product concepts (e.g., Amabile 1998, Weisberg 1993). A firm's professionals "have acquired skills and capabilities that allow them to perform most design tasks more effectively and at a higher level of quality" (Ulrich 2007, Chapter 3, 5ff) and they "often have a significant advantage […] over consumers, in terms of their knowledge, training, and experience" (Moreau and Herd 2010, p. 807).

However, research on the sources of innovation (von Hippel 1988) has long challenged this fundamental assumption and the related, widely held view that users are of little value to firms in providing ideas to solve their unmet consumer problems (as opposed to merely pointing out such problems to firms in the first place). In particular, this line of research has found that many major and minor innovations across various industries were originally developed by users rather than professionals in firms (cf. von Hippel 2005). In their seminal

work, Lilien et al. (2002) provide further evidence that active user involvement in idea generation might benefit a firm's NPD efforts, at least in industrial markets. Specifically, Lilien et al. (2002) compare the potential value of several new industrial product concepts jointly developed by 3M personnel and selected lead users to those developed by more conventional means (i.e., internal developers only). They found that the former outperformed the latter on key innovation indicators that were assessed by 3M managers (e.g., the product concept's novelty compared to the competition or its potential to create an entire new product family). Furthermore, they found that the sales forecasts for concepts developed by lead users were eight times higher than those of internally developed ideas.

One might plausibly argue that working with knowledgeable industrial users is one thing, while working with potential end consumers is another. However, in various consumer goods domains as well, users are frequently found to innovate for themselves, and many of their innovations are commercially attractive (Franke, von Hippel, and Schreier 2006). For example, a recent survey of a representative sample of UK consumers revealed that six percent, or nearly three million consumers, innovated in the domain of household products. Similar national user innovation statistics have been reported for other countries, including the US (5%, or almost 12 million consumers) and Japan (4%, or almost 4 million consumers) (von Hippel, Ogawa, and De Jong 2011).

Thus, user innovation has come of age as many consumer goods firms are challenging the traditional paradigm and experimenting with new ways to more actively integrate users into the idea generation process. In extreme cases, firms like Threadless no longer employ designers but rely exclusively on their user communities to generate new products (Ogawa and Piller 2006). Many established firms, such as LEGO, Dell, and Muji, have followed suit and now complement their in-house efforts with public idea contests known as "crowdsourcing" initiatives, where idea generation is outsourced to a potentially large,

unknown population (the crowd) in the form of an open call (Bayus 2010; Howe 2008; Surowiecki 2004). Thus, crowdsourcing relies on a self-selection process among users who are willing and able to respond to widely broadcast idea generation competitions (Jeppesen and Lakhani 2010; Piller and Walcher 2006; Poetz and Schreier 2012).

Conceptually, the key insight is that some (but, of course, not all) ideas generated by these self-selected users might outperform the ones generated by firm designers (Poetz and Schreier 2012; Schreier, Fuchs, and Dahl 2012). Users may have a competitive edge in idea generation over designers through their experience as consumers. Leading-edge users, in particular, may have tried to solve consumption problems themselves, and their ideas may be commercially attractive because they foreshadow what other consumers will demand in the future (cf. von Hippel 2005). The respective user population that might be activated via crowdsourcing is also naturally much larger and more diverse compared to the team of designers employed by a given firm; a firm's user community may comprise thousands of talented users from highly diverse backgrounds. It follows that the generation of more (and more diverse) ideas increases the odds of generating a few truly exceptional ones (e.g., Gross 1972, Schreier, Fuchs, and Dahl 2012, Surowiecki 2004, Terwiesch and Ulrich 2009).

Empirically, researchers have recently begun to address important questions like what motivates users to participate (e.g., Hertel, Niedner, and Herrmann 2003), how idea contests should be organized (e.g., Boudreau, Lacetera, and Lakhani 2011), how user ideas can be screened effectively (e.g., Toubia and Flores 2007), or how consumers perceive user-driven firms (e.g., Fuchs and Schreier 2011). In a recent study, for example, Schreier, Fuchs, and Dahl (2012) find that consumers associate user-driven firms with higher innovation abilities, that is, they perceive such firms as being better able to generate innovative new products. In a series of experimental studies, they further find that this innovation inference prompts consumers to demand one and the same set of products more strongly. This finding is

especially interesting because objective product properties were kept constant between experimental conditions.

From an NPD perspective, however, the important question is not if the design mode affects *consumer perceptions* of innovation ability, but rather if it affects the *objective* quality of new product ideas actually generated. This question has recently been addressed by Poetz and Schreier (2012), who compared the quality of ideas for new baby products generated by a firm's designers to those generated by users in a public idea contest. Ideas from both sources were judged by company executives (who were blind to the source of ideas), and the key finding was that user ideas were characterized by higher novelty and higher customer benefit, but also by lower feasibility.

Although these findings are promising, it has yet to be explored whether the best of these user ideas, if realized, would eventually perform better on the market. There are several reasons that make this extrapolation non-obvious. First and foremost, it should be noted that in most cases firms still have to translate user ideas into marketable products; a process that demands much effort and involves many decisions that may or may not lead to ultimate market success. More specifically, lower feasibility scores of user ideas point to the possibility that the related promise might never be realized. Second, and on a related note, it could be that high innovativeness and low feasibility eventually boost development and production costs and thereby reduce any potential margin. Third, it is always an empirical question how well *managers' perceptions of ideas* match *consumer reactions to new products* once they reach the shelves. Even if this match is strikingly high (i.e., no judgment and translation errors are involved), it remains unclear how a certain increment in innovativeness (e.g., 10%) will translate into incremental market performance (e.g., X% more sales). Finally, it is worth noting that firms can only market a very limited number of new products at a given point in time. As such, what matters is not the entire distribution and the average idea

quality, but rather the extreme values: the very best ideas available (e.g., Singh and Fleming 2010, Fleming 2007). It remains unclear whether the very best ideas from both sources would differ to such an extent that one could observe managerially meaningful differences in the ultimate consumer reactions.

Against this backdrop, it is surprising that a systematic, empirical market performance assessment of user- vs. designer-generated consumer products is still missing in the literature, despite the broad and growing interest in this topical phenomenon. This is most likely due to the difficulty of obtaining reliable quantitative market data from real-world practice. Such research is important because the market performance of user-driven initiatives will guide marketers in deciding whether to follow the paradigm shift and involve users in generating new product ideas. Drawing on a unique data set gathered from the Japanese consumer goods firm Muji, we address this important gap in the literature. In short, our study reveals that user-generated products systematically outperform designer-generated ones in terms of important outcome variables, including actual sales revenues. Most importantly, we see our study as “initial evidence” in favor of user-generated products and as a contribution to a more critical reflection on the dominant NPD paradigm. As such, this paper makes a second contribution in its conceptual discussion of the generalizability of our findings.

2. Study Method

2.1. New Product Development at Muji

Muji began experimenting with idea contests ten years ago; thus, it is a forerunner among firms that market user-generated products on a broad scale. Muji has also recently marketed a number of user-generated products alongside products generated by the firm’s designers, which allows us to empirically assess the performance of the two approaches. In addition to granting us access to sensitive data (e.g., sales data, gross margins), the firm

provided us with substantial support to complement and validate performance data (e.g., interviews with executives, coding of ideas).

Muji is an established Japanese manufacturer and retailer of a broad range of consumer goods, with a particular focus on interior and household products. Muji's products are sold in almost 500 Muji stores in 22 countries including Japan, China, the US, the UK, France, Italy, and Germany. In 2010, Muji generated approximately 170 billion yen in sales revenue (approximately 2.2 billion dollars), producing net profits of almost 8 billion yen (approximately 100 million dollars). Over the last few decades, Muji has become known for developing and marketing innovative consumer solutions, which is reflected in their receipt of more than 90 design and innovation prizes (e.g., two international Design Oscars awarded by iF International Forum Design, two Good Design Award Gold awards from the Japan Institute of Design Promotion, one Michael E. Porter Prize for innovation). Their ability to generate successful innovations sets a high standard of performance for user-generated products.

For the purposes of our research, we focus on Muji's furniture division. By focusing on one product type, we minimize the potential of confounding effects that arise from comparing different product categories. The furniture division is strategically important to Muji because it accounts for nearly 20 percent of their total sales. We also selected Muji's furniture division because it has experimented more with user-generated products in recent years than other divisions; in other words, data for the furniture division contained the largest number of observations for user-generated products. One might argue that this method of sample selection may bias our results because the furniture division began to rely on user input due to the *under*performance of designers, thereby lowering the performance standard that user-generated products must exceed. In our interviews at Muji, however, we found no evidence for such selection bias.

In 2002, Muji launched the first user-driven piece of furniture, the “Floor Sofa,” which is a large cushion that sits on the floor and adjusts to the body. Because of the novelty and experimental nature of this project, Muji collected binding customer pre-orders before investing in final product development to minimize the risk of failure. The Floor Sofa was a great success. Within the first year, it generated almost 500 million yen in sales, a record that was not exceeded by any of Muji’s designer-generated products until the beginning of our observation period (2005). Prior to 2005, the firm carried out another crowdsourcing experiment that led to the production of a new piece of furniture in 2003 (a small, movable wall shelf) that generated first-year sales of 71 million yen (ranking it sixth among the 15 new furniture products introduced in 2003).

Based on the success of these products, the firm decided to use crowdsourcing on a more systematic basis. During our observation period (February 2005 to July 2009), 43 new products were developed, produced, and introduced to the Japanese market. Thirty-seven products were designer-generated, and six were user-generated. We excluded the two previous user-generated products from our analysis because they differed from the other products in terms of the idea selection process (binding customer pre-orders compared to selection by the firm). However, the results do not change if the observation period is extended to include these two early user-generated products, as well as the respective designer-generated products (results available from the authors upon request). In what follows, we describe the development of designer- and user-generated products in greater detail.

Designer-generated new products. Muji’s conventional approach to NPD adheres closely to the recommendations made in the traditional literature on marketing and innovation. The active agents in idea generation are the firm’s designers, engineers, and marketers, who work in teams organized around specific, new product development projects.

Based on established market research (e.g., market trend studies, focus groups with target consumers), the team identifies and selects important consumption problems to be addressed by a new product (referred to as new product "themes"). In one product development project, for example, the theme was limited storage capacity in consumers' bedrooms due to small room sizes. Based on insight from market research, members of the project team generate new product ideas (e.g., a flat-panel shelf matching the firms' line of beds). The best ideas enter the next phase of product development, where designers propose more detailed design concepts and product specifications based on the product's intended functionality (e.g., a specific design of a flat-panel shelf that is as thin as possible yet large enough to guarantee proper use and stability). The team then decides what concept to move to the next stage. Finally, a new product is introduced to the market. The key characteristic for our study is that it is the internal employees who generate new product ideas aimed to address the selected consumer problem (i.e., the new product theme).

User-generated products. In a second development track, Muji invites users to generate ideas for new products. Anyone who registers on their website can participate. Similar to the internal development of designer-generated products, idea generation follows a specific theme for which solutions can be proposed. Users can also strongly influence the selection of the theme by participating in online discussion boards (as is the case in the designer-generated track). The theme "Suwaru Seikatsu" ("Sit Down Life"), which aimed to generate ideas for better and more comfortable sitting solutions in consumers' homes, was one of the first themes to be successfully outsourced to users. The theme was broadcast to users, who responded with many new product ideas, including the eventual winner: the Floor Sofa. After the first stage of idea generation, users can also comment, vote, and improve on each other's ideas. The best idea is adopted by a Muji project team, which defines the product characteristics in more detail (e.g., external fabrics, cushion filling), seeks further feedback

from the user community, and finally moves the new product to the full production stage. Thus, Muji employees are also heavily involved in the development process for new products in this second track. In contrast to the first track, however, new product ideas are generated by users rather than the firm's designers.

Adopted user ideas in our observation period included the following: a sofa-bed (a sofa that can easily be changed to a bed and vice versa); a wooden cupboard with a movable wagon to increase flexibility of use; a small sofa chair with slim arms to fit in small rooms; a highly versatile stacking shelf; a bed with a thin, yet robust, wooden frame to provide more storage space under the bed; and a sofa with arm rests that are comfortable for both sitting (arms) and lying down (head and legs).

A series of in-depth interviews with senior managers and project members revealed that the two design tracks did not differ substantially on other important dimensions. In particular, projects centered around designer-generated vs. user-generated products did not differ in terms of the human resources invested by the firm (e.g., the quantity and quality of project staff). Furthermore, it is unlikely that employee motivation differed substantially between the two development processes because the firm's employees usually work on several projects at once (including designer- and user-generated products), and both individual and team incentives are identical across projects. Finally, the firm's idea and concept selection criteria are identical in both tracks; a minimum threshold of expected sales and gross margins must be met before a project moves to full production. Interviews with managers further refuted the argument that adopted user-generated ideas "survived a tougher screening" process. Therefore, it is unlikely that any differences in performance between designer- and user-generated products are attributable to these alternative factors.

Nonetheless, the number of ideas generated differed systematically between the two processes, which is consistent with the theoretical work on user involvement and

crowdsourcing (Poetz and Schreier 2012; Schreier, Fuchs, and Dahl 2012). The average estimated number of ideas per theme was approximately ten in the designer condition (this number did not vary much across projects because teams are expected to come up with at least ten good ideas before a choice is made), whereas the *lowest* number of ideas per theme was over 400 in the user condition (the observed maximum was over 2,000). Any random user idea, therefore, has clearly lower odds of being selected, ceteris paribus. However, this does not bias our results. If both distributions empirically ended at the same right-hand point and these top ideas were selected, there should still be no difference in market performance, as there were no differences in the selection criteria employed. If, however, extreme values of the user distribution stretched farther to the right (e.g., because having more ideas increases the likelihood of generating some extreme values; e.g., Gross 1972, Terwiesch and Ulrich 2009), the best user-generated ideas may very well perform better. Such a scenario is consistent with the theoretical arguments that favor user involvement.

2.2. Measures

Independent variable. The independent variable is the source of idea generation: professional designers vs. users (i.e., user-generated vs. designer-generated products).

Dependent variables. Our key outcome variable is the products' actual market performance, measured as aggregate unit sales generated in the first year after market introduction. In addition, we capture profitability by measuring the products' monetary value of sales and gross margins. Because the firm does not allocate fixed costs to individual products, we cannot assess the first-year profits of products, but only their gross margins. Data from the first year that a product is on the market form the most important basis for a product's performance assessment by Muji because furniture products are characterized by short product life cycles (i.e., the first year usually produces the highest sales). To check the

robustness of our results, we also gathered aggregate sales data three years after a product was introduced to the market (the average product life cycle of Muji products is approximately three years), as well as product survival information (i.e., information about which products survived our observation period and were still on the market three years after introduction).

In addition to performance data, we coded all new products (in random order) based on important innovation indicators to gain insight into the nature of the underlying product ideas (e.g., Lilien et al. 2002). First, we asked the general manager of the furniture division and the person in charge of each new product development project to rate the products' novelty relative to competing products and its novelty based on customer needs addressed (measured on a 10-point scale where 1 = low and 10 = high). Because both measures of novelty were highly correlated for both raters ($r_{Rater\ 1} = .52$, $p < .001$, $r_{Rater\ 2} = .84$, $p < .001$), we averaged these measures to create a novelty index. The high correlation between raters ($r_{Between\ raters} = .51$, $p < .001$) suggests that novelty is captured by this index in a valid way (e.g., Franke, von Hippel, and Schreier 2006). We also captured the products' strategic impact on the firm's product portfolio, which we measured using two items: the strategic importance of the product to the business unit's future and the potential of the product to generate a new product line (measured on a 10-point scale where 1 = low and 10 = high). Because both measures were highly correlated for both raters ($r_{Rater\ 1} = .50$, $p < .01$, $r_{Rater\ 2} = .43$, $p < .01$), we averaged them to create a strategic impact index. The high correlation between raters ($r_{Between\ raters} = .47$, $p < .01$) suggests that this variable captures a product's strategic impact in a valid way.

Control variables. We considered the products' price and their year of market introduction as control variables. We re-ran our main analysis (reported below) and controlled for these variables in OLS regressions. As we find that neither price nor inter-year variations

in sales, etc. account for our findings (all reported effects remain significant), we do not report the resulting statistics due to space constraints (results are available upon request).

3. Findings

To compare the market performance of user-generated and designer-generated products, we compared the means of market performance measures for both types of products and tested their statistical significance using t-tests. Given the small sample size of the user design condition (n = 6), we also performed Mann-Whitney U-tests to assess the robustness of our results. The results indicate that both tests lead to the same conclusions (see Table 1A).

---Insert Table 1 about here---

Unit sales. In the first year after introduction, we find that user-generated products sell roughly twice as much as designer-generated products (M_{User} = 30,182.33, $M_{Designer}$ = 14,191.76, $p_{t\text{-}test} < .08$). Similarly, we find that the distribution of the products' first-year unit sales is significantly different between the two conditions ($p_{M\text{-}W\text{-}U} < .05$). Five out of six user-generated products are represented in the top third of the best sellers among all 43 products (ranked 2, 4, 7, 10 and 11), and only one user-generated product had lower unit sales than the average designer-generated product (see Figure 1). The results are even more pronounced using aggregate sales data across three years: units of user-generated products sold three times more frequently than designer-generated ones ($p_{t\text{-}test} < .12$; $p_{M\text{-}W\text{-}U} < .05$).

Sales. First-year sales revenues of user-generated products are more than three times higher than those of designer-generated products (M_{User} = 500.97 million yen, $M_{Designer}$ = 141.04 million yen, $p_{t\text{-}test} < .05$). Similarly, the distribution of sales revenues differs significantly between the two types of products ($p_{M\text{-}W\text{-}U} = .001$). Four of the top six products, in terms of sales, are user-generated (ranked 1, 2, 3 and 6), and the two remaining user-generated products generated higher sales revenues than the average designer-generated

product (greater than 147.80 million yen, see Figure 2). Aggregate sales data across three years suggests that the size of this effect increases over time: three years after introduction, user-generated products generated an average of 1,250 million yen (approximately 16 million dollars) *more* in sales than designer-generated products, which is equal to over five times the sales of designer-generated products ($p_{t\text{-}test} < .05$; $p_{M\text{-}W\text{-}U} = .001$).

Gross margin. User-generated products yielded gross margins (M_{User} = 240.47 million yen) that were approximately four times higher than designer-generated products ($M_{Designer}$ = 56.22 million yen, $p_{t\text{-}test} < .05$). The distribution of gross margins is also significantly different between the two conditions ($p_{M\text{-}W\text{-}U} = .001$). In particular, four of the top five products, in terms of gross margins, are user-generated products (ranked 1, 2, 3 and 5) and the remaining two user-generated products yielded higher gross margins than the average designer-generated product (greater than 67.90 million yen; see Figure 3). The results are more pronounced if we use three years of sales data: the average margin of user-generated products is 619 million yen (approximately 8 million dollars) *higher*, or six times greater, than designer-generated products ($p_{t\text{-}test} < .05$; $p_{M\text{-}W\text{-}U} = .001$). These findings suggest that user-generated products demonstrate stronger market performance than their designer-generated counterparts.

Because performance should be related to survival, we additionally analyzed how many designer- and user-generated products survived our observation period (three years after introduction). Products that survived performed substantially better than those which were discontinued (e.g., first-year unit sales were approximately 60% higher for products that survived the third year). Indeed, only 17 out of 37 designer-generated products were still on the shelves after three years. In contrast, five out of six user-generated products were still on the market ($p_{chi\text{-}square\text{-}difference\text{-}test} < .09$). User-generated products, therefore, have a substantially higher survival rate at Muji.

---Insert Figure 1 through 3 about here---

Novelty and strategic impact. The results with regard to the nature of each new product provide corroborating evidence for the findings reported above. In particular, we find that user-generated products are characterized by substantially higher novelty compared to designer-generated products (M_{User} = 7.29, $M_{Designer}$ = 4.78, $p_{t\text{-}test} < .001$, $p_{M\text{-}W\text{-}U} < .001$). Four of the top five products, in terms of novelty, are user-generated (ranked 1, 2, 3 and 4) and the remaining two user-generated products are rated among the top 50% of the sample on novelty and consistently exhibit higher novelty than the average designer-generated product (greater than 5.90, see Figure 4). The results for the products' strategic impact are similar: user-generated products have a significantly more pronounced strategic impact on the firm's future product portfolio than designer-generated ones (M_{User} = 6.08, $M_{Designer}$ = 4.25, $p_{t\text{-}test}$ = .001, $p_{M\text{-}W\text{-}U} < .01$). Three of the top five products, in terms of strategic impact, are user-generated (ranked 1, 3 and 4), and the "worst" user-generated product has a strategic impact score (4.00) close to the average score of designer-generated products (see Figure 5).

---Insert Figure 4 and 5 about here---

Sensitivity analysis. Although we only analyzed products within a single category (furniture), the sample consists of different products (e.g., sofa, table, chair, bed, shelf). To assess whether different product types within the furniture category might partially account for the findings reported above (beyond price considerations), we re-analyzed the data using a more homogeneous subsample of products. We matched the six user-generated products with similar designer-generated ones (n = 15; sofa, bed, shelf). As Table 1B shows, the results remain robust; user-generated products outperform designer-generated ones on all dimensions.

In a second robustness check, we aimed to broaden the scope of our analysis. In particular, we sought market performance data from another category, namely health and

beauty products. This category has brought about four user-generated products in recent years (in the area of natural and homeopathic remedies and skin care) where the respective product launch was far enough in the past in order to track aggregate performance data across three years. Table 2 shows that user-generated products once again demonstrate better market performance. Although we identified several other Muji categories with user-generated products (e.g., kitchenware, stationery, gardening products, and clothing), user-generated products in these categories were either too low in number (e.g., one or two products) and/or market introduction had occurred too recently (e.g., in 2010).

---Insert Table 2 about here---

4. Discussion

There is broad and growing interest in understanding the topical phenomenon of more actively involving users in firms' idea generation processes. Yet and despite conceptual advancements (Poetz and Schreier 2012; Schreier, Fuchs, and Dahl 2012) it is surprising that an empirical market performance assessment of user-driven products (i.e., products that are based on user- vs. designer-generated ideas) is still missing in the literature. This is most likely due to the difficulty of obtaining reliable market data from real-world practice. As such, it remains unclear whether it will pay off for firms to depart from the classic NPD paradigm anchored in the idea that firm professionals – and not users – should generate the ideas for new products to be marketed to broader customer segments. By drawing on a unique data set gathered from the Japanese consumer goods firm Muji, which has drawn on both sources of ideas in parallel in recent years, we provide initial research to fill this gap in the literature. We demonstrate that user-generated products systematically and substantially outperform their designer-generated counterparts in terms of key market performance metrics.

Of course, we do not see our study as the “final battle” that would once and for all relegate the “loser” to the history books. Rather, we see our study as initial evidence and as a contribution to a more critical reflection on the dominant NPD paradigm. In a nutshell, our findings suggest that it might pay for firms to draw on user ideas *in parallel* to their established in-house efforts – not unlike what Muji seems to have identified as their best practice. Furthermore, users were only the providers of ideas; the overall NPD process entails many more stages, and it goes without saying that decisive in-house efforts and capabilities are needed to convert any promising idea into a successful new product. Bearing this in mind, we now discuss the generalizability of our findings and identify promising directions for future research.

First, it is worth noting that the majority of winning user-designers in our sample were highly involved in the category, had enduring past exposure to the problems for which ideas were sought, had relatively high levels of technical expertise, and thus generally matched the conceptual description of lead users (cf. von Hippel 2005). Clearly, if firms are unable to activate this type of user, it is unlikely that the effects found in this study can be replicated. The notion of users’ leading-edge status is also conceptually important. Different types of users might account for the discrepancy between predictions made in the classical marketing and NPD literatures, on the one hand, and research on user innovation, on the other hand. Whereas the former might have “average” users in mind (which would justify skepticism about user involvement because these users lack both technical and usage expertise), the latter typically constructs their arguments around leading-edge users. In other words, conventional marketing is concerned with the “average user” and research on user innovation targets only users with leading-edge status who represent “extreme values” on the right-hand side of the distribution of users.

The consequence of involving the “right” vs. “wrong” users might loom even larger for more complex products than the ones studied in this paper (Poetz and Schreier 2012; Schreier, Fuchs, and Dahl 2012). For complex products (e.g., cars), the expertise required for successful ideation might be too high, thereby preventing the vast majority of average users from contributing meaningful ideas. On a related note, all of the furniture themes studied were based on a clear “usage problem” (e.g., improving the sitting experience in the living room). We argue that such problems are well-suited for users because users have a competitive edge over designers in terms of the needs-based information necessary to solve the problem effectively. If, in contrast, the theme was based on technical details regarding technologies or materials, company designers would be more likely than users to have the skills and expertise to solve these problems. These speculations offer promising avenues for future research.

Furthermore, the openness of contests may favor user-based ideas. The user population that is accessed through crowdsourcing is naturally much larger and more diverse compared with a team of designers employed by a given firm. It follows that, if more (and more diverse) ideas are generated, the odds of generating a few truly exceptional ideas will increase (e.g., Gross 1972, Schreier, Fuchs, and Dahl 2012, Surowiecki 2004, Terwiesch and Ulrich 2009). Our data from Muji support this prediction: out of the many ideas submitted by users (ranging from 400 to more than 2,000 per theme), the best user ideas outperformed those generated by designers. Yet the relationship between the openness of contests, how the contest announcement is targeted, the activation of effective self-selection mechanisms, and the final outcome is not well understood; thus, it constitutes fruitful ground for future research.

Interviews with managers at Muji further revealed that they have succeeded in building the specific capability to interact with users effectively. Nonetheless, each crowdsourcing

initiative took substantial time to implement, is generally less controllable and manageable in the early stages, and it cannot easily be implemented on a large scale (as reflected in the relatively small number of user-generated compared with designer-generated new products). Thus, our findings might not be easily replicated with data from “old-school” firms with no prior experience in active user involvement. As a result, one important, but unanswered question remains: what are the specific capabilities that make user-driven firms successful?

Finally, it is worth noting that Muji did not actively market user products as “designed by users” to the general public. We, therefore, only assessed the “objective” market performance of user-driven products. As indicated above, however, recent research suggests that firms might benefit from a “designed by users” label over and above any objective product characteristics (Schreier, Fuchs, and Dahl 2012). To assess the full potential of marketing user-generated products, future research should examine this psychological effect on performance.

In conclusion, we note that further research is required to fully understand the newly emerging paradigm of an active user involvement in NPD. Our study provides important initial evidence that such efforts might pay off.

References

Amabile, Teresa M. (1998), "How to Kill Creativity," *Harvard Business Review* 76(5), 76-87.

Bayus, Barry L. (2010), "Crowdsourcing and Individual Creativity Over Time: The Detrimental Effects of Past Success," working paper (August), (http://papers.ssrn.com).

Boudreau, Kevin J., Nicola Lacetera and Karim R. Lakhani (2011), "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis," *Management Science*, in print.

Cooper, Robert G. (2001), *Winning at New Products: Accelerating the Process from Idea to Launch* (3rd edition), Cambridge, MA: Perseus Books.

Crawford, C. Merle and C. Anthony Di Benedetto (2000), *New Products Management* (6th edition), Burr Ridge, IL: Irwin-McGraw Hill.

Fleming, Lee (2007), "Breakthroughs and the 'Long Tail' of Innovation," *Sloan Management Review*, Fall 49(1), 69-74.

Franke, Nikolaus, Eric von Hippel, and Martin Schreier (2006) ,"Finding Commercially Attractive User Innovations: A Test of Lead User Theory," *Journal of Product Innovation Management*, 23(4), 301-15.

Fuchs, Christoph and Martin Schreier (2011), "Customer Empowerment in New Product Development," *Journal of Product Innovation Management*, 28(1), 17-32.

Gross, Irwin (1972), The Creative Aspects of Advertising, S*loan Management Review*, Fall 14(1), 83-109.

Hertel, Guido, Sven Niedner and Stefanie Herrmann (2003), "Motivation of Software Developers in Open Source Projects: An Internet-based Survey of Contributors to the Linux Kernel," *Research Policy*, 32(7), 1159–77.

Howe, Jeff (2008), *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business*, New York: Crown Business.

Jeppesen, Lars Bo and Karim R. Lakhani (2010), "Marginality and Problem-Solving Effectiveness in Broad Search," *Organization Science*, 21(5), 1016-33.

Lilien, Gary L., Pamela D. Morrison, Kathleen K. Searls, Mary Sonnack and Eric von Hippel (2002), "Performance Assessment of the Lead User Idea Generation Process," *Management Science*, 48(8), 1041-59.

Moreau, C. Page and Kelly B. Herd (2010), "To Each His Own? How Comparisons with Others Influence Consumers' Evaluations of Their Self-Designed Products," *Journal of Consumer Research*, 36(February), 806-19.

Ogawa, Susumu and Frank T. Piller (2006), "Reducing the Risks of New Product Development," *MIT Sloan Management Review*, 47(2), 65-71.

Piller, Frank T. and Dominik Walcher (2006), "Toolkits for Idea Competitions: A Novel Method to Integrate Users in New Product Development," *R&D Management*, 36(3), 307-18.

Poetz, Marion K. and Martin Schreier (2012), "The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas?," *Journal of Product Innovation Management*, 29(2),245-256.

Schreier, Martin, Christoph Fuchs, and Darren Dahl (2012), "The Innovation Effect of User Design: Exploring Consumers' Innovation Perceptions of Firms Selling Products Designed by Users," *Journal of Marketing*, 76(September), 18-32.

Schulze, Anja and Martin Hoegl (2008), "Organizational Knowledge Creation and the Generation of New Product Ideas: A Behavioral Approach," *Research Policy*, 37(10), 1742-750.

Singh, Jasjit, and Lee Fleming (2010), "Lone Inventors as Sources of Technological Breakthroughs: Myth or Reality?," *Management Science*, 56(1) ,41-56.

Surowiecki, James (2004), *The Wisdom of Crowds – Why the Many Are Smarter Than the*

Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations, New York: Doubleday.

Terwiesch, Christian and Karl T. Ulrich (2009), *Innovation Tournaments: Creating and Selecting Exceptional Opportunities*, Boston: Harvard Business Press.

Toubia, Olivier, and Laurent Flores (2007), "Adaptive Idea Screening Using Consumers," *Marketing Science*, 26(3), 342–60.

Ulrich, Karl T. (2007), *Design: Creation of Artefacts in Society*, Pontifica Press.

Ulrich, Karl T. and Steven D. Eppinger (1995), *Product Design and Development* (1st edition), New Jersey: McGraw-Hill.

Urban, Glen L. and John R. Hauser (1993), *Design and Marketing of New Products* (2nd edition), Englewood Cliffs, NJ: Prentice Hall.

Von Hippel, Eric (1988), *The Sources of Innovation*, Oxford University Press, New York.

Von Hippel, Eric (2005), *Democratizing Innovation*, MIT Press, Boston.

Von Hippel, Eric, Susumu Ogawa and Jeroen P.J. de Jong (2011), "The Age of the Consumer-Innovator," *MIT Sloan Management Review*, 53(1), 27-35.

Weisberg, Robert W. (1993), *Creativity: Beyond the Myth of Genius*, New York: Freeman

Table 1: Summary of Findings (Mean Statistics)

(A) Analysis of the full sample (furniture division)

	Designer-generated products (n = 37)	User-generated products (n = 6)	
	M	M	Significance
Unit sales (after 1st year)	14,191.76	30,182.33	$p_{M\text{-}W\text{-}U} < .05$; $p_{t\text{-}test} < .08$
Unit sales (after 3rd year)	29,959.65	91,120.50	$p_{M\text{-}W\text{-}U} < .05$; $p_{t\text{-}test} < .12$
Sales in yen (after 1st year)	141.04 million	500.97 million	$p_{M\text{-}W\text{-}U} = .001$; $p_{t\text{-}test} < .05$
Sales in yen (after 3rd year)	276.46 million	1,525.63 million	$p_{M\text{-}W\text{-}U} = .001$; $p_{t\text{-}test} < .05$
Gross margin in yen (after 1st year)	56.22 million	240.47 million	$p_{M\text{-}W\text{-}U} = .001$; $p_{t\text{-}test} < .05$
Gross margin in yen (after 3rd year)	116.52 million	735.69 million	$p_{M\text{-}W\text{-}U} = .001$; $p_{t\text{-}test} < .05$
Newness[1]	4.78	7.29	$p_{M\text{-}W\text{-}U} < .001$; $p_{t\text{-}test} < .001$
Strategic impact[1]	4.25	6.08	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} = .001$

(B) Sensitivity analysis (product-type matching of user- and designer-generated products)

	Designer-generated products (n = 15)	User-generated products (n = 6)	
	M	M	Significance
Unit sales (after 1st year)	8,905.93	30,182.33	$p_{M\text{-}W\text{-}U} < .05$; $p_{t\text{-}test} < .08$
Unit sales (after 3rd year)	18,305.53	91,720.50	$p_{M\text{-}W\text{-}U} < .05$; $p_{t\text{-}test} < .08$
Sales in yen (after 1st year)	174.43 million	500.97 million	$p_{M\text{-}W\text{-}U} < .05$; $p_{t\text{-}test} < .07$
Sales in yen (after 3rd year)	313.03 million	1,525.63 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .05$
Gross margin in yen (after 1st year)	69.23 million	240.47 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .06$
Gross margin in yen (after 3rd year)	127.81 million	735.69 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .05$
Newness[1]	4.93	7.29	$p_{M\text{-}W\text{-}U} < .001$; $p_{t\text{-}test} < .001$
Strategic impact[1]	4.47	6.08	$p_{M\text{-}W\text{-}U} < .001$; $p_{t\text{-}test} < .01$

[1]Measured on a 10-point scale where 1 = low and 10 = high

Table 2: Sensitivity Analysis (Health and Beauty Division)

(A) Analysis of the full sample			
	Designer-generated products (n = 49)	User-generated products (n = 4)	
	M	M	Significance
Unit sales (after 1st year)	45,527.76	74,787.50	$p_{M\text{-}W\text{-}U} < .07$; $p_{t\text{-}test} < .25$
Unit sales (after 3rd year)	106,452.61	170,995.25	$p_{M\text{-}W\text{-}U} < .13$; $p_{t\text{-}test} < .38$
Sales in yen (after 1st year)	13.00 million	69.10 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .09$
Sales in yen (after 3rd year)	28.63 million	162.64 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .06$
Gross margin in yen (after 1st year)	4.38 million	23.50 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .09$
Gross margin in yen (after 3rd year)	9.97 million	57.50 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .06$
(B) Sensitivity analysis (product-type matching of user- and designer-generated products)			
	Designer-generated products (n = 12)	User-generated products (n = 4)	
	M	M	Significance
Unit sales (after 1st year)	38,428.50	74,787.50	$p_{M\text{-}W\text{-}U} < .09$; $p_{t\text{-}test} < .15$
Unit sales (after 3rd year)	74,878.50	170,995.25	$p_{M\text{-}W\text{-}U} < .06$; $p_{t\text{-}test} < .13$
Sales in yen (after 1st year)	7.20 million	69.10 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .08$
Sales in yen (after 3rd year)	12.14 million	162.64 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .05$
Gross margin in yen (after 1st year)	2.05 million	23.50 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .07$
Gross margin in yen (after 3rd year)	3.99 million	57.50 million	$p_{M\text{-}W\text{-}U} < .01$; $p_{t\text{-}test} < .05$

Note that we also analyzed survival in this category. Importantly, we again find that only 21 out of 49 designer-generated products were still on the shelves after three years. At the same time, however, all four user-generated products were still on the market ($p_{chi\text{-}square\text{-}difference\text{-}test} < .05$). User-generated products thus demonstrate a substantially higher survival rate. (Products that survived the three-year observation period also performed substantially better than those that were discontinued, which confirms that lower-performing products were eliminated; e.g., first-year unit sales were approximately 140% higher for products that survived the third year.)

Figure 1: Unit Sales per Product (Furniture)

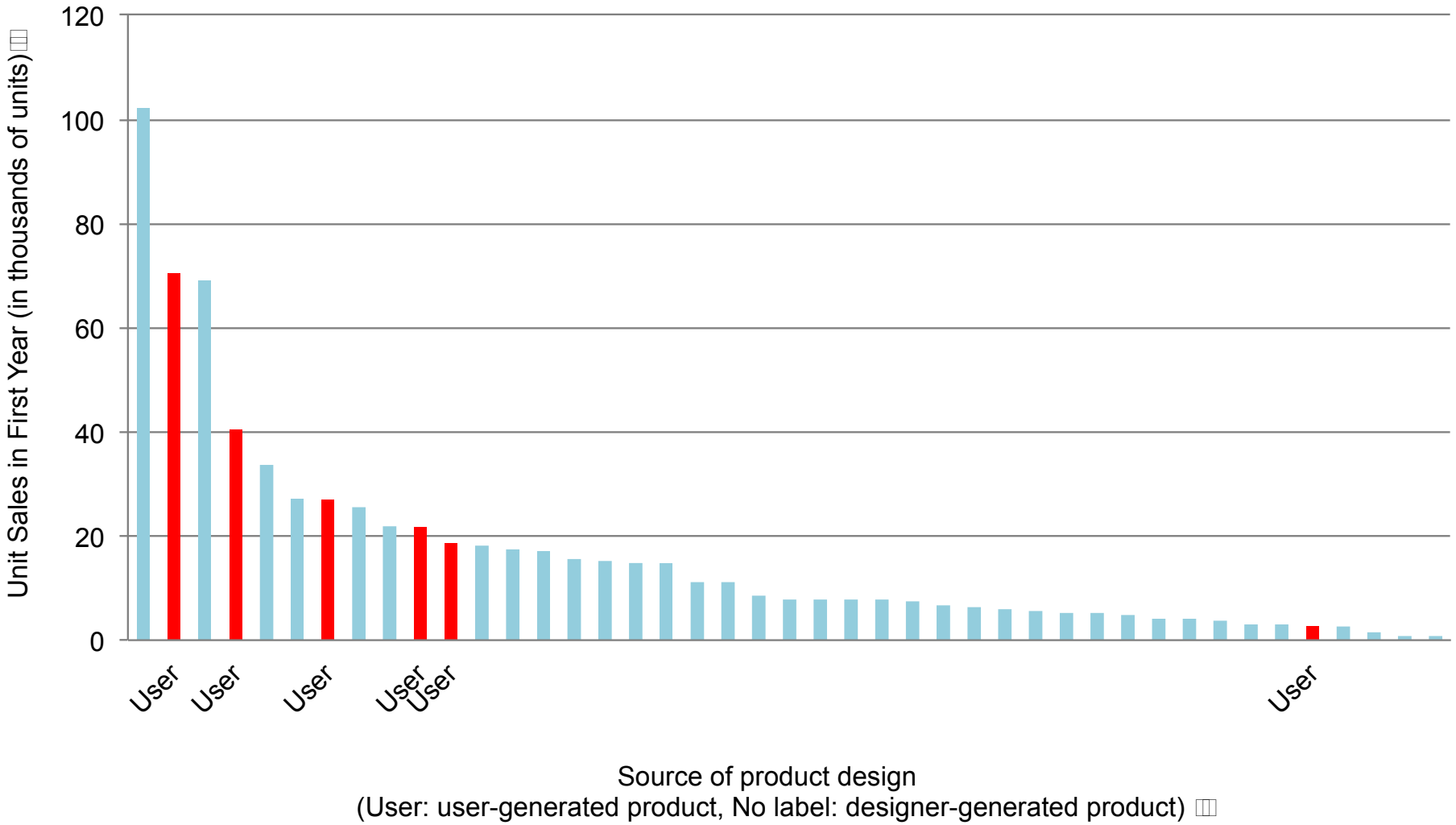

Figure 2: Sales per Product (Furniture)

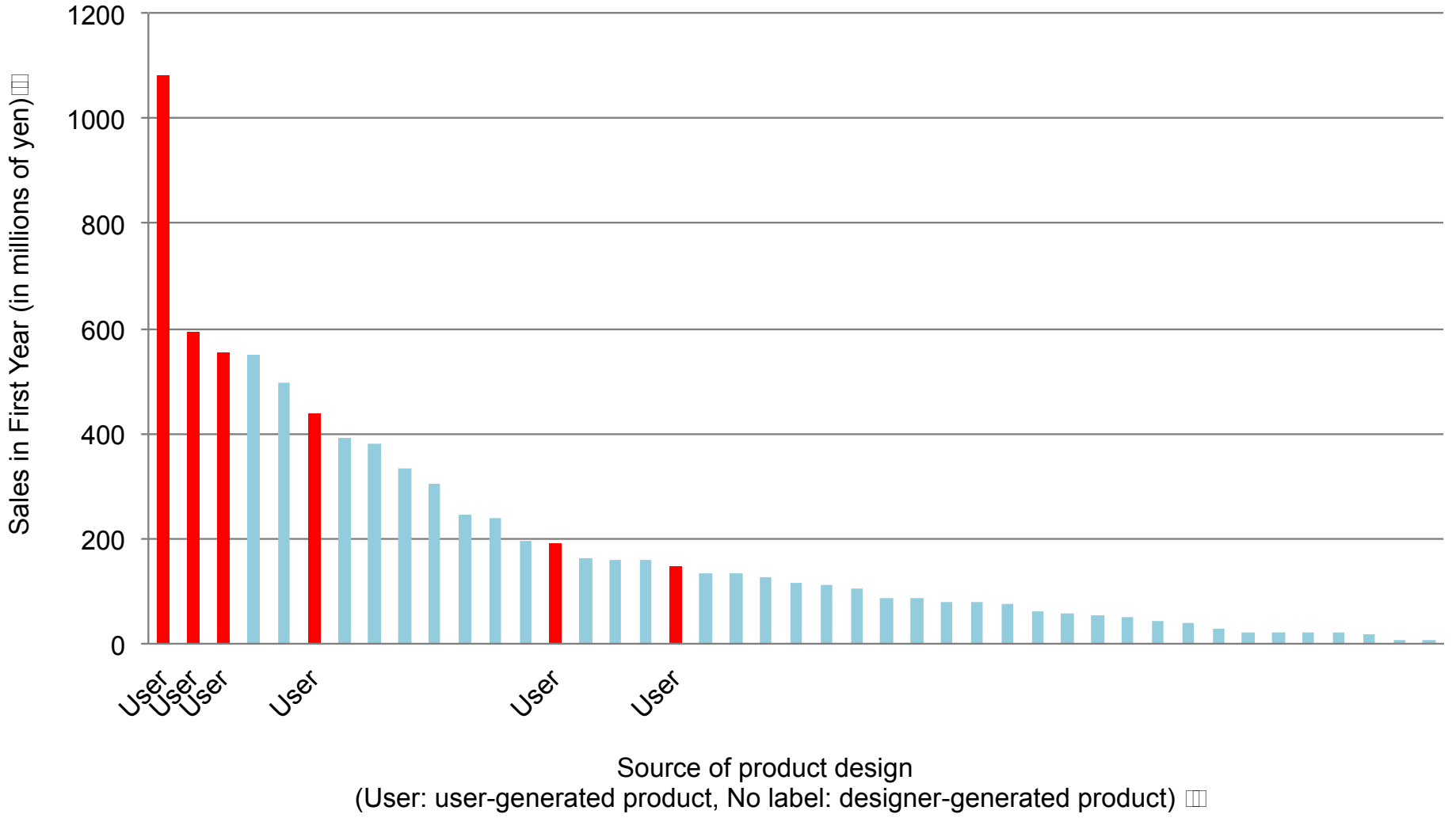

Figure 3: Gross Margin per Product (Furniture)

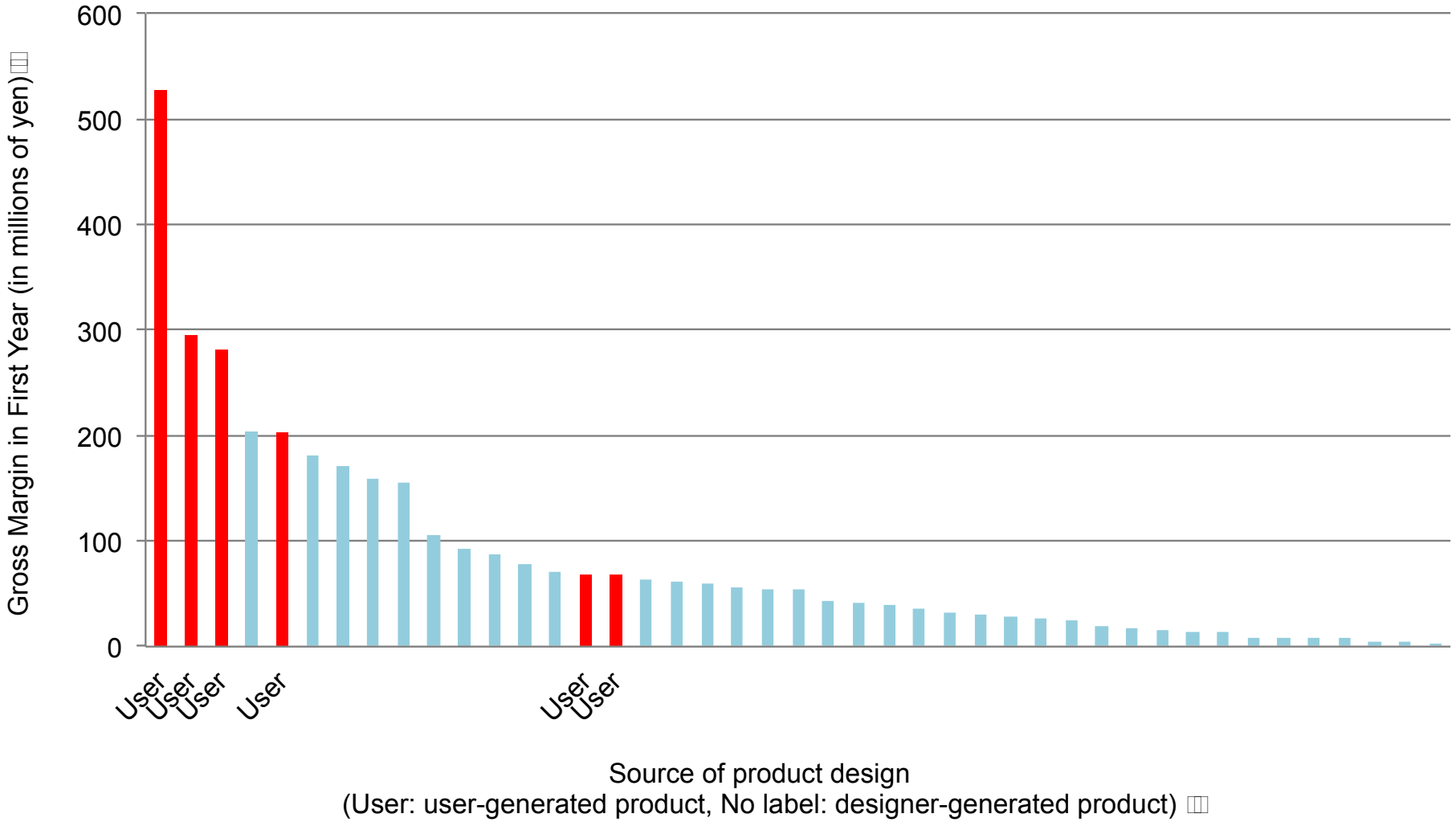

Figure 4: Newness per Product (Furniture)

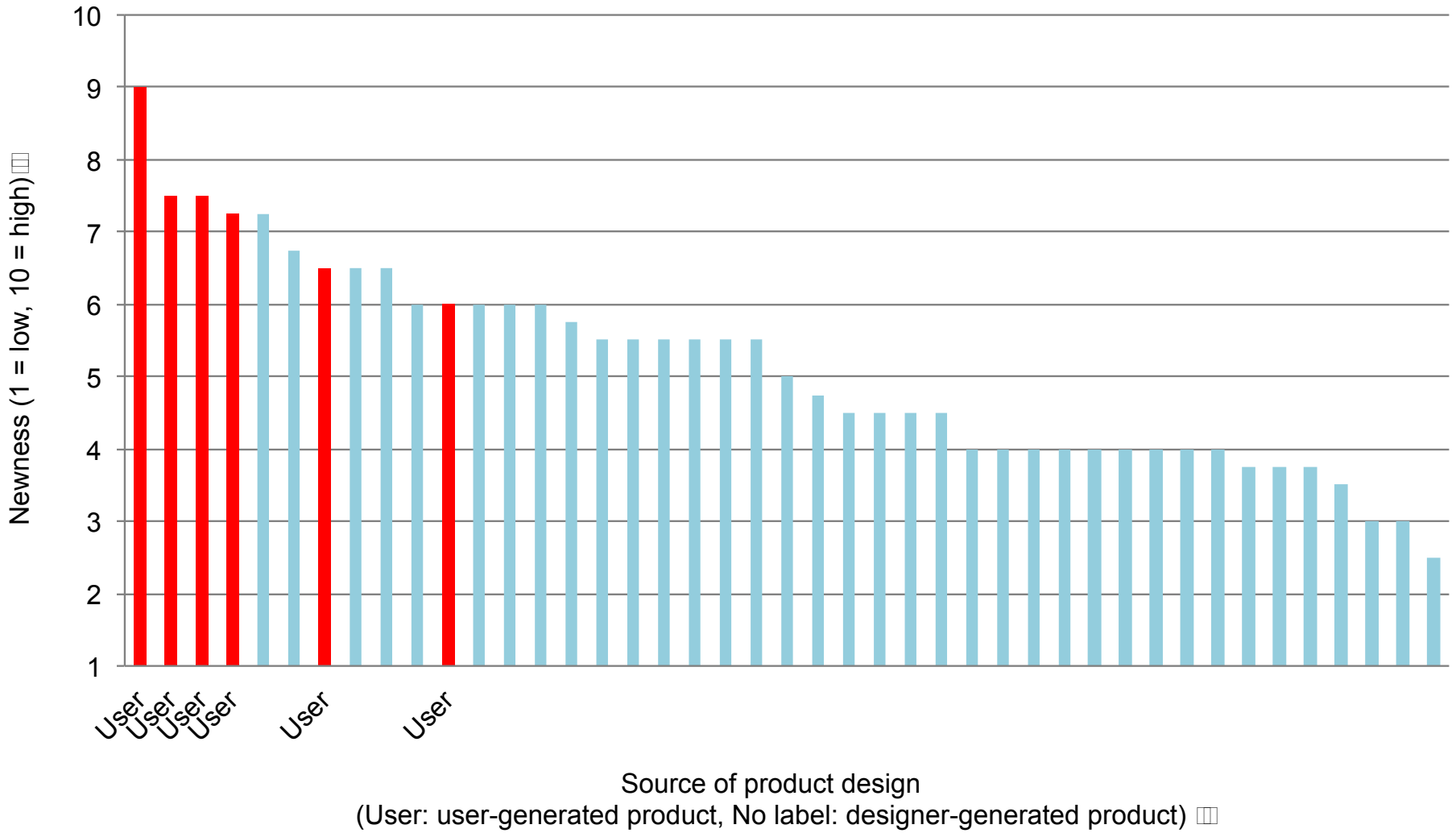

Figure 5: Strategic Impact per Product (Furniture)