# Can Consumers Learn Price Dispersion?

# Evidence for Dispersion Spillover Across Categories

WORKING PAPER – DO NOT CITE OR DISTRIBUTE WITH THE AUTHORS' APPROVAL

Quentin André

Nicholas Reinholtz

Bart de Langhe

Quentin André (corresponding author; andre@rsm.nl) is an assistant professor of marketing at the Rotterdam School of Management, Erasmus University. Nicholas Reinholtz (nicholas.reinholtz@colorado.edu) is an assistant professor of marketing at the University of Colorado Boulder. Bart de Langhe (bart.delanghe@esade.edu) is an associate professor of marketing at Universitat Ramon Llull, ESADE. The authors wish to thank Dan Goldstein, and their colleagues at the University of Colorado Boulder, ESADE, INSEAD, and Rotterdam School of Management, for their comments and feedback. This paper is based on Quentin André's dissertation, for which Nicholas Reinholtz and Bart de Langhe served as advisors.

## ABSTRACT

Dispersion knowledge—the beliefs that consumers have about the minimum, the maximum, and the overall variability of values in a distribution—is a key antecedent of many consumer judgments and decisions. This paper examines consumer's ability to form accurate dispersion knowledge in multi-category environments. Ten studies provide evidence for a phenomenon we call *dispersion spillover*: Seeing more (vs. less) dispersed prices in one category inflates people's perception of price dispersion in another category. Dispersion spillover is consequential: It influences judgments of price attractiveness, the likelihood that consumers will search for (and find) better options, and how much consumers bid in auctions. In several studies, we examine the mental representations that underlie dispersion spillover with a graphical "distribution builder" tool. Our data suggest that dispersion knowledge is not only based on exemplar memory: People also seem to form and apply "intuitive statistics" that summarize the distributions they have seen, but these representations of dispersion are not entirely category-specific. To encourage and facilitate future investigations, we share distBuilder, an open-source JavaScript library that allows researchers to add distribution builders to their studies with minimal programming knowledge and effort. (186 words)

*Keywords*: consumer learning, numerical cognition, intuitive statistics, price search, price knowledge, reference price, price dispersion, mental representation.

## INTRODUCTION

Consumers often make decisions based on quantitative data (e.g., prices, calories, miles, quality ratings, time), and effective decision making requires them to consider the distributional properties of this data. While knowledge about the central tendency—or average value—of a distribution may be sufficient for some decisions, optimal decision-making often necessitates a more nuanced understanding of the variability—or spread—of a distribution. Such "dispersion knowledge" is often acquired through experience. For instance, by repeatedly shopping for groceries, a consumer will over time learn about the price distribution for breakfast cereals. However, the multiplicity of categories in consumers' environment complicates this learning. A simple trip to the grocery store will expose a consumer to prices from dozens of product categories. Navigating this complexity—and effectively learning through experience—requires consumers to form category-specific dispersion knowledge (e.g., the dispersion of prices for breakfast cereals, the dispersion of prices for juices, and so on). This paper examines the formation of dispersion knowledge in multi-category environments and how dispersion knowledge influences consumer decisions.

## DISPERSION KNOWLEDGE IS IMPORTANT

Imagine you are planning to purchase a new car. You have decided on a make, model, and trim, and are now contacting dealers to get price quotes. To make a purchase decision, it is helpful to know the average price for this specific type of car across dealers. This would allow you to determine if a given quote is "cheap" or "expensive." However, what is more helpful still is to know the minimum price you could potentially pay, or how likely it is that another dealer's price quote will be better than the quotes you have received so far. Knowledge about the dispersion of a distribution is critical to weigh the costs and benefits of information search. This problem of price search, and the analogous problem of wage search, have been studied extensively (McCall 1972; Rothschild 1974; Stigler 1962). The automobile search problem described above is actually a motivating example from early work (Stigler 1961), but the same idea holds true for consumer search problems more generally. A core result from these models is that the utility of search depends on the amount of price dispersion, and that consumers should search more when there is more dispersion (Reinholtz 2015).

Beyond consumer search, knowledge of dispersion is valuable in countless other contexts. Consider consumer financial decision making: A core tenant of investing is to maximize expected returns given an acceptable level of risk, or to minimize risk given a desired level of expected returns, where "risk" is typically operationalized as variability in returns (Markowitz 1952). Similarly, a common prescription in household finance is to have an emergency savings fund. Knowledge about one's average income and expenses is not sufficient to make a decision about the size of an emergency fund. This should be determined by the frequency and magnitude of unexpected expenses (Gallagher et al. 2018; Skinner 1988).

Additionally, consider some other commonplace consumer decisions: You have an important meeting tomorrow at 8:00am. What time should you leave your home? If your average morning commute time is 50 minutes, would you feel comfortable leaving at 7:10am? Suppose you found the home of your dreams in a hot housing market where the owner is going to take the best offer. Her realtor tells you the average of the other bids is $600,000. How likely would you be to get the house with an offer of $625,000?

Knowledge of the average value is insufficient in these—and many other—contexts. To make informed decisions, consumers need to have accurate beliefs about the dispersion of values. The expected frequency and magnitude of "cheap" prices when searching for a product, the extremity of price fluctuations of an asset when making an investment, the likelihood and size of unexpected expenses when saving for emergencies, the longest it might reasonably take when planning a commute, and the most likely values of the highest bids when competing in an auction. We contend that the formation—and use—of these beliefs is a critical question in the study of consumer behavior.

## CONTRIBUTIONS

This paper aims to make three contributions. First, in a series of incentive-compatible experiments, we demonstrate that consumers can make poor decisions about products and services in one category because of the prices they have seen in other categories. In particular, when price dispersion in other categories is high, consumers are more likely to reject attractive offers, search for better prices that do not actually exist, and bid excessively in auctions.

Second, we provide evidence that these effects are driven by a distortion in dispersion knowledge. Seeing more (vs. less) dispersed prices in one category leads people to perceive more price dispersion in another category. Consumers' impressions of dispersion are thus not independent across categories, a phenomenon we call *dispersion spillover*.

Third, we use distribution builders (a graphical elicitation tool that allows people to describe their perception of a numerical distribution by allocating a number of "balls" to "buckets") to gain novel insights about the mental representations that underlie dispersion spillover. We show that a simple misattribution process—according to which people would accurately remember prices but confuse the corresponding category (i.e., remembering the price of a white wine as the price of a red wine)—cannot fully explain the dispersion spillover. Instead, our results suggest that the dispersion spillover reflects the formation and application of *intuitive statistics of dispersion* across categories. Consumers appear to rely on summary representations of the amount of dispersion that they have encountered, but those abstract representations are not entirely category-specific. To facilitate future investigations of dispersion knowledge, we have developed distBuilder, an open-source JavaScript library that allows academics and business practitioners to add distribution builders to their surveys with minimal programming knowledge and effort.

## THEORETICAL BACKGROUND

The present work is connected to past research—which we summarize below—that has examined how people perceive statistical properties (central tendency, dispersion, correlation, etc.) and how well their perceptions correspond to the true statistical properties of a distribution (Gigerenzer and Murray 2015; de Langhe 2016; Peterson and Beach 1967).

### Knowledge of Central Tendency

The earliest investigations in this domain focused on people's ability to learn the central tendency of a distribution (mean, median, or mode). In the typical paradigm, participants see a list of numbers and are asked to report one or several measures of central tendency. These studies revealed remarkably accurate estimates (e.g., Levin 1974; Spencer 1963), although people sometimes conflated mean and median (Peterson and Miller 1964; Winkler 1970), and tended to make more mistakes when confronted with large values (Levin 1975), a larger number of values (Smith and Price 2010), and values with greater dispersion (Laestadius 1970; Spencer 1963; Wolfe 1975).

Studies also revealed that people are well-equipped to learn the central tendencies of multiple categories presented concurrently. When presented with quantitative information from two categories, people were able to accurately report the average value for one category without being influenced by the information from the other category (Levin 1974, 1975; Malmi and Samson 1983). The ability to form category-specific estimates of central tendency is essential for many types of judgments and decisions. For instance, it allows consumers to form a reference price for multiple brands (Kalwani et al. 1990) and to develop expectations of how much money they will earn in different circumstances (Crawford and Meng 2011).

**Knowledge of Dispersion**

Decades of literature on cue learning (Brehmer 1973; Brehmer and Lindberg 1970; Mellers 1986; Mellers, Richards, and Birnbaum 1992), categorization (Fried and Holyoak 1984; Rips 1989; Sakamoto, Jones, and Love 2008; Stewart and Chater 2002), signal detection (Verghese 2001; Wilken and Ma 2004), and risky choice (March 1996; Weber, Shafir, and Blais 2004) imply that people's knowledge of dispersion informs their judgments and decisions. However, we know relatively little about people's ability to form dispersion knowledge from experience, and in particular, about their ability to learn dispersion in multi-category environments.

The few empirical investigations that exist are difficult to interpret. In a first type of paradigm, participants were presented with a sequence of numbers and then asked to report “the variance” or “the standard deviation” (Lovie 1978; Lovie and Lovie 1976). Participants’ estimates were far from the true values, but it is not clear what to conclude from this observation. Even financial analysts, who probably should understand the statistical definition of “standard deviation”, fail to provide accurate numerical estimates of standard deviation for financial products (Goldstein and Taleb 2007). Statistical definitions of dispersion may be especially unintuitive for lay people, and poor performance may reflect a problem of translation rather than a lack of knowledge.

In a second type of paradigm, participants were presented with stimuli from two categories (with different underlying distributions) and then asked to report the ratio of the variances (Beach and Scopp 1968) of those categories, or simply to indicate which category had higher variance (Kareev, Arnon, and Horwitz-Zeliger 2002). However, for studies using such a comparative approach, it is again not entirely clear what can be learned from poor performance. Does it reflect that people acquire little dispersion knowledge through experience, that their intuitive understanding of dispersion is different from common statistical benchmarks, or that they conflate the dispersions of two distributions?

**Distribution Builders**

To shed light on consumers' knowledge of distributions, some researchers have recently relied for a graphical elicitation technique called "distribution builders." Instead of directly eliciting estimates of summary statistics, the tool asks participants to create a histogram representing their perception of the entire distribution, by allocating a fixed number of "balls" (representing frequencies) to "buckets" that represent possible values or ranges of values.

Originally developed to aid in retirement savings decisions (Goldstein, Johnson, and Sharpe 2008), distribution builders have since been used to measure investors' beliefs about stock returns (Long, Fernbach, and De Langhe 2018; Reinholtz, Fernbach, and De Langhe 2016), to probe preferences for the distribution of wealth within a population (Page and Goldstein 2016), and to test people's preferences for risky outcomes (Dietvorst and Bharti 2019). More generally, Goldstein and Rothschild (2014) have demonstrated that people describe an observed distribution more accurately using a distribution builder compared to more-traditional, non-graphical methods (e.g., stated fractiles, stated quantiles, or confidence intervals). They therefore advocate for the use of distribution builders to elicit knowledge of quantitative information.

Despite all their advantages, distribution builders are not currently available in popular survey engines (e.g., Qualtrics) and therefore cannot be implemented without significant programming knowledge and effort. To help address this issue and facilitate the adoption of distribution builders, we have developed distBuilder, an open-source JavaScript library that allows academics and business practitioners to use distribution builders with minimal programming knowledge and effort. A link to the source code and documentation for this library is available at the following address: [link removed to keep peer-review blind].

## SUMMARY OF STUDIES

This paper presents 10 studies. In the typical study paradigm, participants undergo a learning phase where they observe many prices from two different categories of products. By manipulating the means and variances of the price distributions in the learning phase, we can examine people's ability to learn the properties of the two distributions, and the extent to which properties of one distribution influence people's judgments and decisions for the other distribution. The studies are presented in three parts.

Part 1 (studies 1–3) provides evidence for *dispersion spillover*: People ascribe more price dispersion to a distribution when another distribution that they learned concurrently had a large (vs. small) amount of price dispersion. For instance, participants in study 1 were more likely to report seeing a white wine priced at $13—while the cheapest white wine they saw was in fact priced at $17—when the price range for red wines was wide versus narrow.

Part 2 (studies 4–7) examines the *downstream consequences* of dispersion spillover. Because of dispersion spillover, participants in study 4 judged cheap prices as less attractive, participants in study 5 spent more money on airfare for an upcoming business trip, participants in study 6 rejected attractive bonus payments, and participants in study 7 made excessive bids in an auction for a gift card.

Part 3 (studies 8–10) uses the distribution builder methodology to examine the *mental representations* that underlie dispersion spillover. Our data suggest that dispersion spillover cannot be explained based on an imperfect storage and retrieval of previously seen prices. Instead, our data suggests that people form abstract mental representations ("intuitive statistics") to summarize distributions, and that these abstract representations are not entirely category-specific.

# PART 1: DISPERSION SPILLOVER

## Study 1: Dispersion Spillover for Minimum and Maximum Values

Study 1 features a learning phase, in which participants are exposed to prices from two categories, and a test phase, in which participants answer questions related to the distributions of prices they saw in the learning phase. Prices in one category were the same for all participants. We refer to it as the "common" category. Prices in the other category were manipulated between participants. We refer to it as the "manipulated" category. The price dispersion in the manipulated category was either the same as in the common category, larger, or smaller. In the test phase, we asked participants to report the minimum and maximum prices they remembered seeing in each of the two categories (i.e., common category and manipulated category). We then examined how the amount of actual price dispersion in the manipulated category affected participants' judgements of price dispersion for the manipulated category and for the common category.

### *Method*

We paid 300 respondents from Amazon Mechanical Turk (MTurk) 50 cents for their participation. We excluded data from seven participants because the maximum price they reported seeing was strictly lower than the minimum price they reported seeing for at least one of the two categories, leaving us with 293 responses. We did not record any demographic or psychographic variables.

In the learning phase, all participants saw 26 prices for red wines and 26 prices for white wines. All prices were presented in a random order, such that the prices for reds and whites were intermixed. We presented each price for 1.2 seconds, together with a picture of a bottle of red or white wine. We instructed participants to pay close attention to the type of wine corresponding to each price because we would ask questions about the prices of each type of wine later in the study. To facilitate learning, we asked participants to say the prices aloud as they viewed them.

Red and white wines had the same average price ($25), but we manipulated the dispersion of prices between participants. The common category had the same, medium amount of price dispersion across conditions (with prices ranging between $17 and $34; SD = 4.5). The manipulated category had a small (with prices ranging between $23 and $28; SD = 1.1), medium, or large (with prices ranging between $13 and $38; SD = 7.5) amount of price dispersion.[1] To avoid the possibility that prior beliefs about prices of red and white wines would introduce a bias, we counterbalanced between participants which type of wine (red vs. white) was assigned to each category. There were no significant effects of this counterbalancing factor, so we do not elaborate on it further.

In the test phase, we measured participants' perceptions of dispersion by asking them to report the "most expensive" and the "cheapest" wine they saw in each category (i.e., cheapest white, most expensive white, cheapest red, and most expensive red).

### *Results*

If consumers can form accurate dispersion knowledge, we should observe two patterns. First, participants should report more extreme minimum and maximum prices for the

[1] The exact price distributions presented in each study are reported in the online appendix.

manipulated category when the prices presented in this manipulated category had a large (vs. medium vs. small) amount of price dispersion. Second, participants should report similar minimum and maximum prices for the common category across conditions.

Figure 1 shows the average minimum (top panel) and maximum (bottom panel) prices participants reported for the manipulated (left) and the common category (right). For the manipulated category, participants appropriately reported seeing less extreme minimum prices when the actual price dispersion in the manipulated category was low ($M_{low}$ = 20.42) versus medium ($M_{med}$ = 18.36) versus high ($M_{high}$ = 15.03; all ps < .001). They also appropriately reported seeing less extreme maximum prices when the actual price dispersion in the manipulated category was low ($M_{low}$ = 30.44) versus medium ($M_{med}$ = 33.46) versus high ($M_{high}$ = 35.45; all ps < .001).

For the common category, the minimum and maximum prices participants report should not depend on the price dispersion in the manipulated category. Participants appropriately reported similar minimum prices when the price dispersion in the manipulated category was medium versus low ($M_{med}$ = 17.88 vs. $M_{low}$ = 18.19: $p$ = .444), and they appropriately reported seeing similar maximum prices when the price dispersion in the manipulated category was medium versus low ($M_{med}$ = 32.65 vs. $M_{low}$ = 33.18: $p$ > .371). However, participants *inappropriately* reported seeing more extreme minimum prices when the price dispersion in the manipulated category was high versus medium ($M_{high}$ = 16.23 vs. $M_{med}$ = 17.88: z = -3.99, p < .001, standardized $b$ = -.50), and they *inappropriately* reported seeing more extreme maximum prices when the price dispersion in the manipulated category was high versus medium ($M_{high}$ = 34.57 vs. $M_{med}$ = 32.65: z = 3.19, $p$ < .001, standardized $b$ = .43). So, when the manipulated category had greater price dispersion than the common category, participants seem to have assimilated this greater dispersion into their beliefs about the dispersion of the common category.

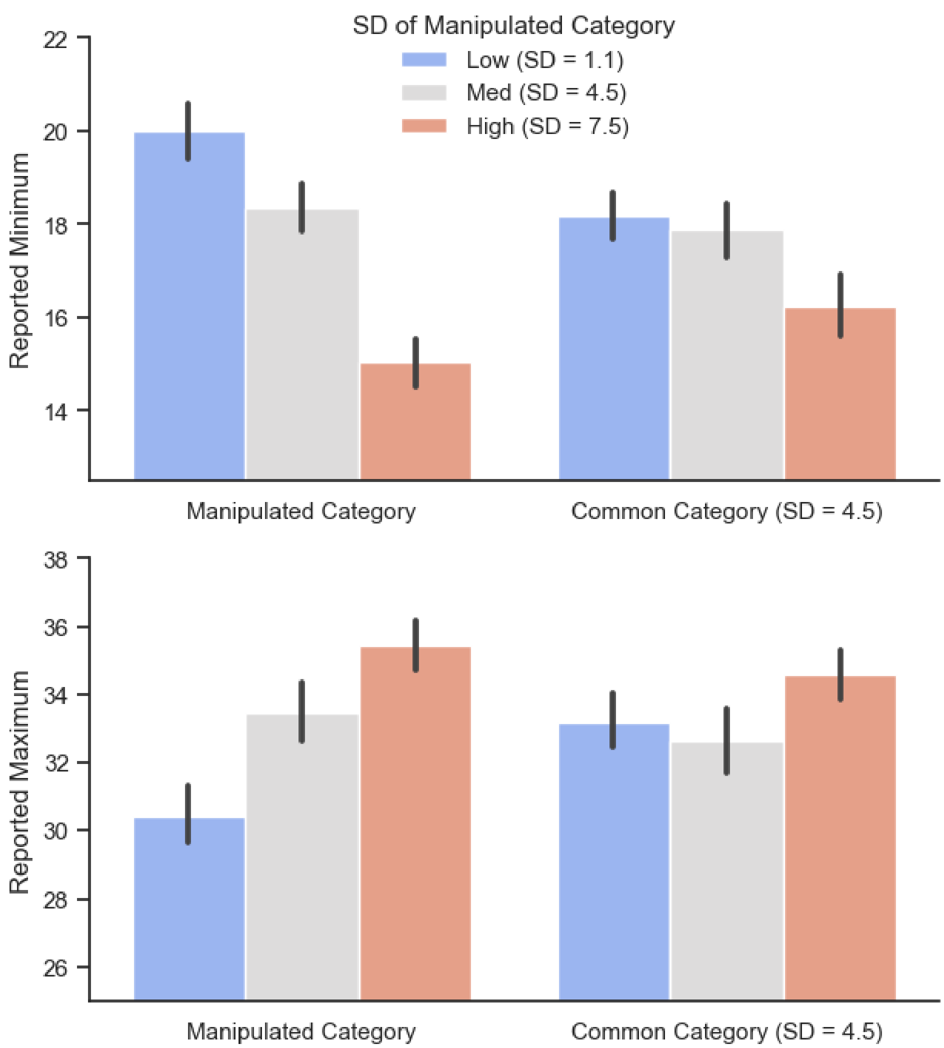


Figure 1. Reported minimum (top panel) and maximum (bottom panel) prices in study 1. Error bars are 95% bootstrapped confidence intervals.

Figure 2 shows the cumulative distributions for the minimum (Figure 2a) and maximum (Figure 2b) prices that participants reported for the common category. When the manipulated category had less or as much price dispersion, many participants were able to accurately report the true minimum price of the common category. We observe a sharp discontinuity at $17, and few participants reported a minimum price smaller than this value (same: 12%, lower: 9%; $\chi(1) = 0.41$, $p = 0.520$). The same is true of the maximum price. We observe a discontinuity at $34, and few participants reported a maximum price larger than this value (same: 17%, lower: 16%, $\chi(1) = 0.06$, $p = 0.799$). In contrast, when the manipulated category had a larger amount of price dispersion, 53% of participants underestimated the true minimum price of the common category (53% vs. 12%: $\chi(1) = 42.90$, $p < .001$, d = 1.06) and 49% of participants overestimated the true maximum price of the common category (49% vs. 16%: $\chi(1) = 22.46$, $p < .001$, d = 0.72).

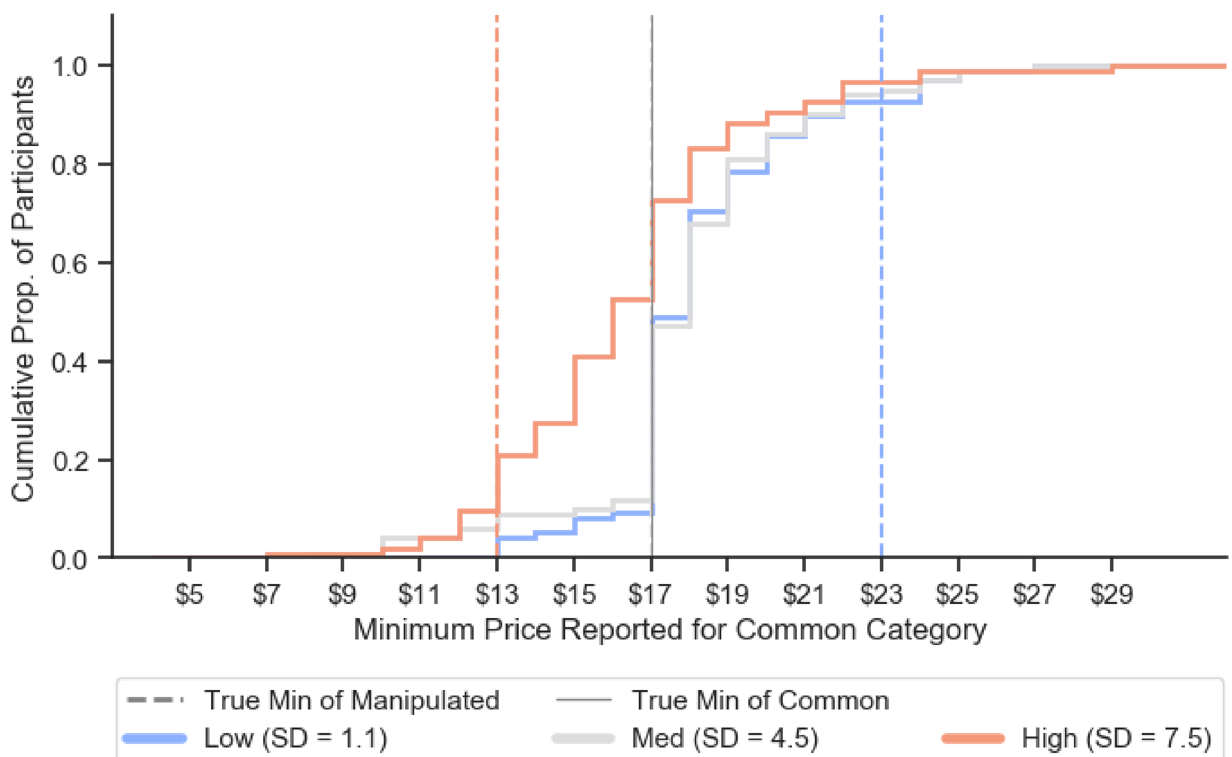


Figure 2a. Cumulative distributions of the reported minimum for the common category in study 1, split by conditions.

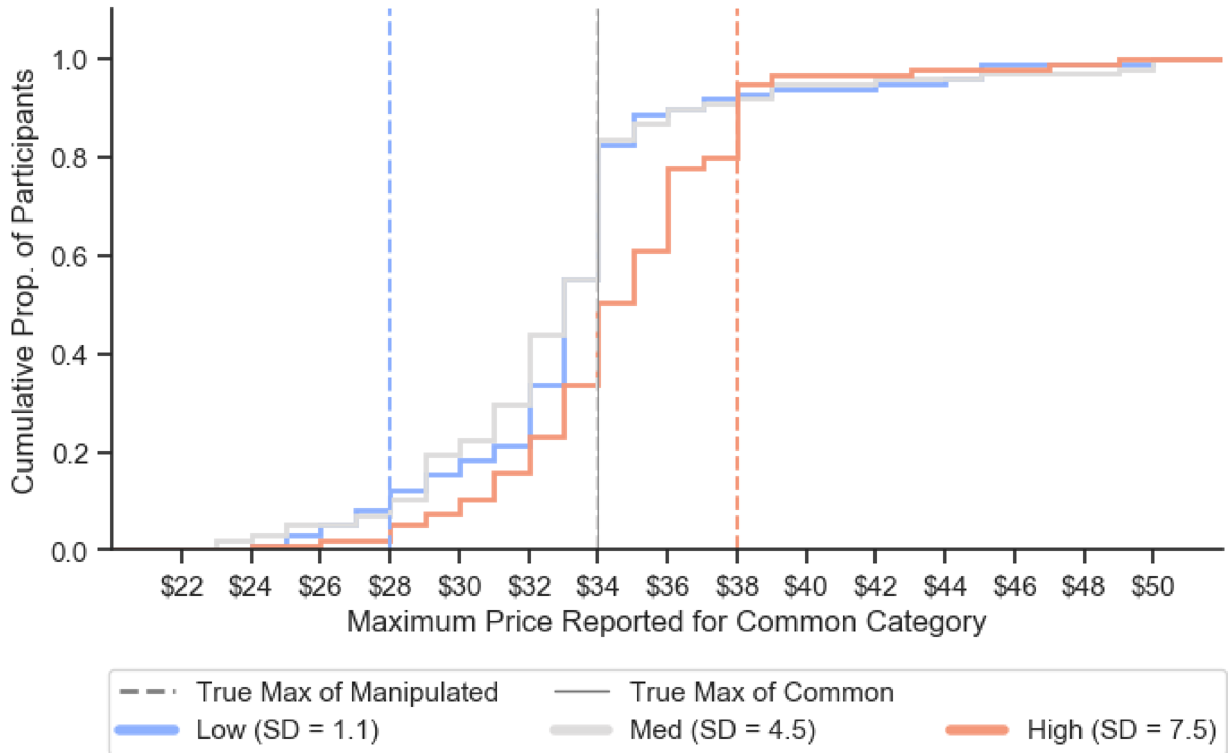


Figure 2b. Cumulative distributions of the reported maximum for the common category in study 1, split by conditions.

In sum, study 1 shows a *dispersion spillover* across categories. Participants' impressions of dispersion for the common category (measured by the perceived minimum and maximum price) were significantly influenced by the amount of dispersion in the manipulated category. This dispersion spillover is asymmetric. People perceived greater price dispersion in a category when another category they learned concurrently had more dispersed prices, but they did not perceive less price dispersion when the other category had less dispersed prices. We return to, and discuss, this asymmetry in part 3.

## Study 2: Dispersion Spillover Using Distribution Builders

Participants in study 1 were presented with many prices, each price was presented only briefly, and the prices came from two similar product categories. These characteristics of the learning phase may complicate learning. In study 2, we modify the learning phase in two ways to examine the robustness of the dispersion spillover. First, instead of seeing one price every 1.2 seconds, we allowed participants to review the prices at their own pace. Second, instead of learning the prices of two types of wines, participants learned the prices of pillows and blankets. We also made a change to the test phase. Participants in study 1 reported the minimum and maximum price they remembered seeing in each category. While the notions of minimum and maximum price are intuitive to participants, they imperfectly capture the amount of dispersion of a distribution. In study 2, we use the distribution builder tool to measure participants' perceptions of the two distributions.

### *Method*

We paid 300 respondents from MTurk 70 cents for their participation. Due to a technical glitch, we were unable to record data from four respondents, leaving a final sample size of 296. We did not record any demographic or psychographic variables.

The design of the learning phase was identical to that of the previous study, except that participants saw prices for pillows and blankets, and that they controlled the presentation speed. Participants could view the prices at their own pace, and they took on average 93 seconds to review all prices, which is 50% longer than in study 1. The product category labels were again counterbalanced across the common and manipulated categories. We did not observe a significant effect of this counterbalancing factor and do not elaborate on it further.

After the learning phase, we asked participants to produce two distributions, one for pillows and one for blankets, on two separate distribution builders. For each category, participants were instructed to construct a histogram of the prices they remembered seeing by allocating 26 markers (one for each pillow/blanket presented in the learning phase) to 25 buckets (corresponding to prices from $1 to $49, in increments of $2).[2] Figure 3 shows an example of a distribution constructed on the distBuilder interface.

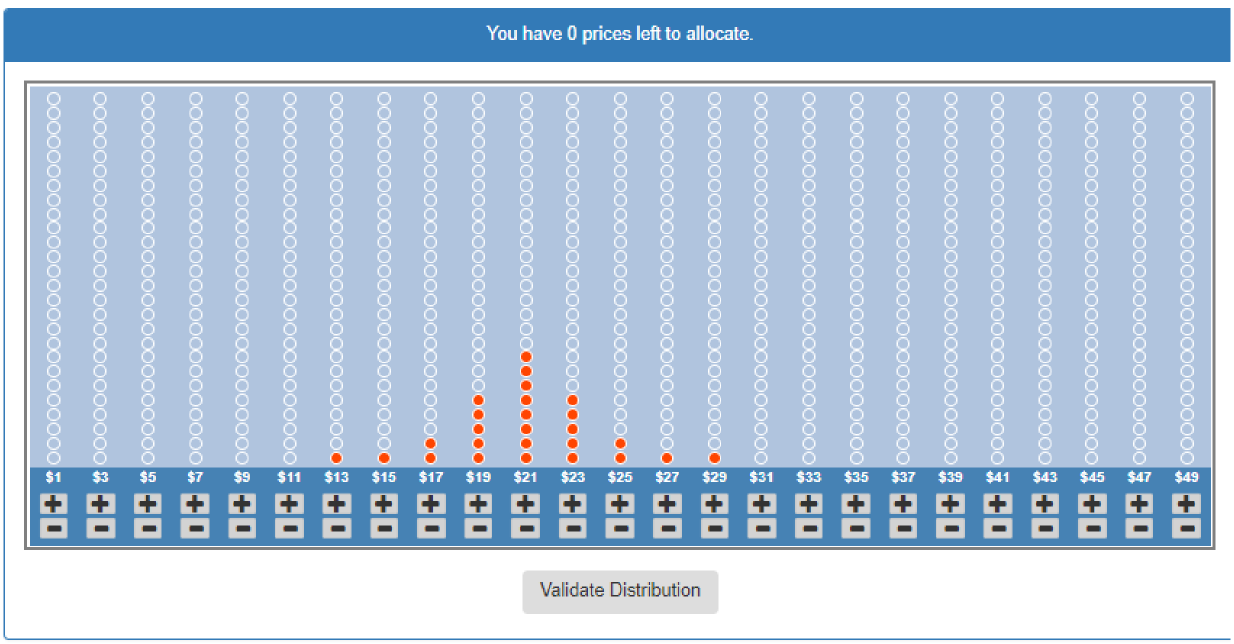


Figure 3. A distribution of prices reported on the distBuilder interface.

### *Results*

[2] A demo of the experimental procedure used in study 2 is available here: [https://dispersion-spillover-preview.herokuapp.com/logout](https://dispersion-spillover-preview.herokuapp.com/logout).

Each participant created two distributions in the test phase, one for the manipulated category, and one for the common category. For each distribution created by participants, we computed the standard deviation of the 26 values reported. Figure 4 shows the average standard deviations of the price distributions, by condition and category. Participants created appropriately wider (tighter) distributions for the manipulated category when the actual price dispersion in this category was higher (lower). However, replicating the dispersion spillover observed in study 1, participants inappropriately created wider distributions for the common category when price dispersion in the manipulated category was high (vs. medium or low).

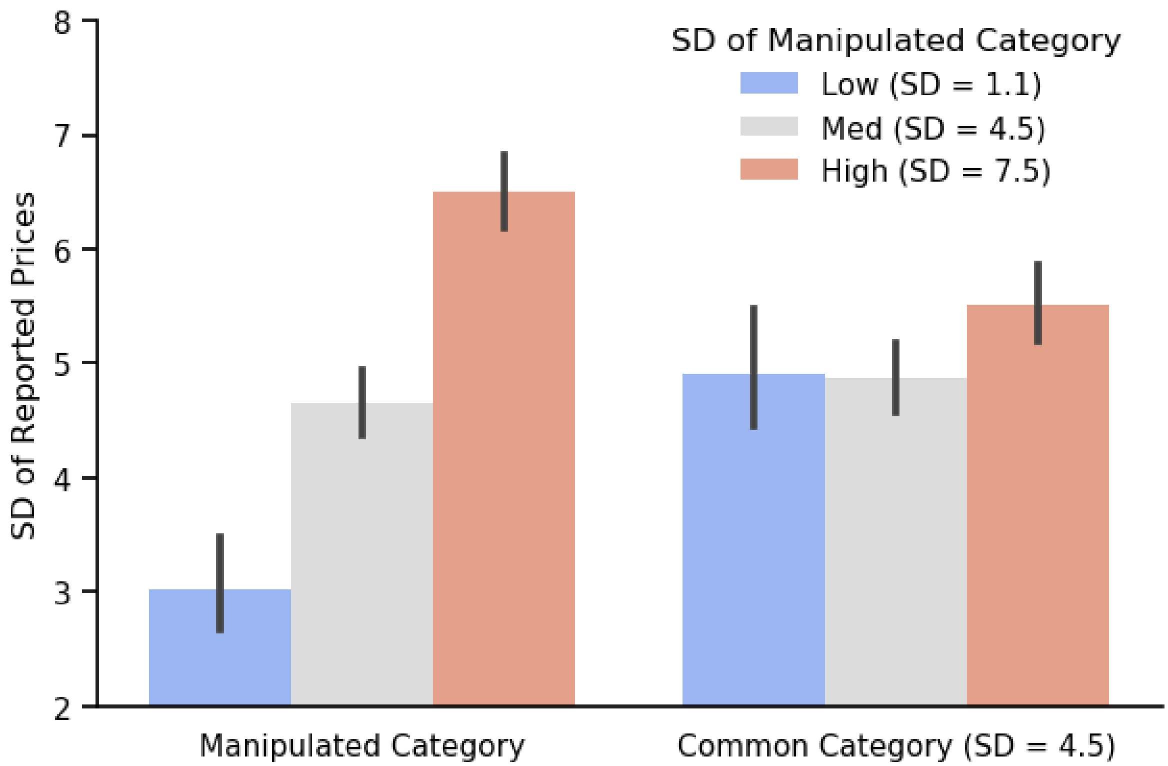


Figure 4. Average standard deviation of distributions created by participants in study 2. Error bars are 95% bootstrapped confidence intervals.

To examine statistical significance, we estimated a mixed linear model with category type (common vs. manipulated) as a within-participant factor, the actual price dispersion of the manipulated category (low vs. medium vs. high) as a between-participant factor, and the interaction between these two factors. For the manipulated category, participants created wider distributions if the price dispersion in the manipulated category was high versus medium versus low ($M_{\text{high}}$ = 6.51 vs. $M_{\text{med}}$ = 4.65 vs. $M_{\text{low}}$ = 3.03, $p$ < .001 for all pairwise comparisons, standardized $b$ for smallest difference = .75). For the common category, participants created wider distributions if the price dispersion in the manipulated category was high versus medium ($M_{\text{high}}$ = 5.52 vs. $M_{\text{med}}$ = 4.87: z = 2.39, $p$ = .017, standardized $b$ = .30), but they created distributions with similar standard deviations if the price dispersion in the manipulated category was medium versus low ($M_{\text{med}}$ = 4.87 vs. $M_{\text{low}}$ = 4.90, $p$ = .897). Thus, we again find an asymmetry, such that a category with a higher price dispersion inflates perceptions of dispersion for a category with a lower price dispersion, but not *vice versa*.

One advantage of using the distribution builder tool is that we can operationalize perceived dispersion in various ways, not just the standard deviation, and that we can observe other properties of the distribution, such as its central tendency. In the online appendix, we present analyses for various measures of dispersion (i.e., variance, minimum, maximum, IQR, and range) and central tendency (i.e., mean, median, and mode). The dispersion spillover observed on the standard deviation is consistent across different measures of dispersion. For instance, participants created a price range for the common category that was $2.61 wider when the price dispersion in the manipulated category was high versus medium (z = 3.14, $p$ < .001, standardized $b$ = .39), while the price ranges they created were similar when the price dispersion in the manipulated category was medium versus low (z = 0.16, $p$ = .869). In contrast, we do not find that manipulating the dispersion had any discernable impact on people's impressions of central tendencies. The online appendix describes these additional analyses for all studies using distribution builders.

In sum, study 2 suggests that the dispersion spillover is robust when people can learn the prices at their own pace, and shows the value of using distribution builders to study consumers' perceptions of dispersion.

## Study 3: Dispersion Spillover when Categories are Learned Sequentially

Participants in studies 1 and 2 learned about the prices from two categories simultaneously, with prices of both categories presented intermixed in a random order. This characteristic of the learning phase may complicate learning. To further examine the robustness of dispersion spillover, participants in study 3 learned about prices from two categories sequentially, one category at a time.

### *Method*

We paid 300 respondents from MTurk 50 cents for their participation. Due to a technical glitch, we were unable to record data from ten respondents, leaving a final sample size of 290. We did not record any demographic or psychographic variables. The learning phase was the same as in study 1, except that the prices in both categories were now presented sequentially. For example, participants might first see prices for 26 bottles of red wine and then see prices for 26 bottles of white wine. We counterbalanced the presentation order of the categories and, as in study 1, we counterbalanced which type of wine (red vs. white) was assigned to each category.

There were no significant effects of these counterbalancing factors, so we do not elaborate on them further. The test phase was the same as in study 2.

***Results***

Each participant created two distributions in the test phase, one for the manipulated category, and one for the common category. As in study 2, we computed the standard deviation of the 26 values reported for each of the two categories. We then estimated a mixed linear model with category type (common vs. manipulated) as a within-participant factor, the actual price dispersion of the manipulated category (low vs. medium vs. high) as a between-participant factor, and all two-way interactions between these two factors.

For the manipulated category, participants created wider distributions if the price dispersion in the manipulated category was high versus medium versus low ($M_{\text{high}}$ = 6.56 vs. $M_{\text{med}}$ = 5.03 vs. $M_{\text{low}}$ = 3.02, $p$ < .001 for all pairwise comparisons, standardized $b$ for smallest difference = .85). For the common category, participants created wider distributions if the price dispersion in the manipulated category was high versus medium ($M_{\text{high}}$ = 5.71 vs. $M_{\text{med}}$ = 4.95: z = 2.41, $p$ = .016, standardized $b$ = .81), but they created distributions with similar standard deviations if the price dispersion in the manipulated category was medium versus low ($M_{\text{med}}$ = 4.95 vs. $M_{\text{low}}$ = 4.35, $p$ = .507). Thus, replicating the previous studies, we again find an asymmetric dispersion spillover.

In sum, this study suggests the dispersion spillover is also observed in circumstances when consumers learn about the price distributions one category at a time.

## PART 2: DOWNSTREAM CONSEQUENCES

## Study 4: Judgments of Price Attractiveness

In study 4, we explore downstream consequences of dispersion spillover. Participants learned about flight prices to Colorado and Florida. We expect that when flights prices to one destination have a large amount of price dispersion, consumers will overestimate the price dispersion of flights to the *other* destination, thus expecting cheaper flights than the ones that were actually presented. As a consequence, they will find offers for cheap flights less attractive, and report being more likely to search for better prices.

### *Method*

We paid 301 respondents from MTurk 60 cents for their participation. We did not record any demographic or psychographic variables. No participants were excluded.

In the learning phase, all participants saw 26 prices for flights to Colorado and 26 prices for flights to Florida, in random order, each presented for 1.2 seconds. To facilitate learning, we displayed Colorado flights in dark blue together with a picture of a mountain, and Florida flights in orange together with a picture of a beach.

Colorado and Florida flights had the same average price ($320), but we manipulated the dispersion of prices between participants. The common category had the same amount of price dispersion across conditions (with prices ranging between $240 and $400; SD = 40). The manipulated category either had the same amount of dispersion as the common category or a higher amount of dispersion (with prices ranging between $140 and $500; SD = 96). As in previous studies, we counterbalanced the labels (i.e., destinations) across categories. We did not observe a significant effect of this counterbalancing factor and do not elaborate on it further.

In the test phase, we asked participants questions about the common category (i.e., Colorado or Florida, depending on the counterbalancing factor). Thus, participants answered questions about a category that had an identical amount of price dispersion across conditions. First, we asked participants to report the cheapest flight they remembered seeing. If there is dispersion spillover, people should report seeing cheaper flight prices when the manipulated category had greater (vs. equivalent) dispersion. Second, we presented participants with five offers of flights priced at \$280, \$260, \$240, \$220 and \$200. Two offers were more expensive than the minimum price presented to participants in the common category (\$280 and \$260), one was equal to the minimum price (\$240), and two were cheaper than the minimum price (\$220 and \$200). We presented the offers sequentially, in decreasing order, on separate pages, and asked participants to indicate for each offer how likely they would be to accept it (vs. search for a better price) on a scale from –3 (“definitely prefer searching for a better price”) to +3 (“definitely prefer accepting the deal”). If there is dispersion spillover, they should believe that cheaper prices exist in the category, and should find these offers less attractive.

***Results***

As in previous studies, we found a dispersion spillover such that the dispersion of flight prices in the manipulated category affected participants’ beliefs about the cheapest flight in the common category. Figure 5 shows the cumulative distribution of minimum prices reported by participants. The actual cheapest flight was \$240. When the dispersion of the manipulated category was equal to that of the common category, there is a sharp discontinuity at this value, indicating that people formed an accurate perception of the minimum price. Only 10% of participants reported seeing a flight cheaper than \$240, and the average minimum price reported by participants was \$250. In contrast, this discontinuity at the true minimum price is far less pronounced when the dispersion of the manipulated category was higher: 58% of participants reported seeing a flight cheaper than the minimum price (58% vs. 10%: $\chi(1) = 77.92$, $p < .001$, d = 1.18), and the average minimum price reported by participants was \$211 (\$211 vs. \$240: $t(299) = -7.13$, $p < .001$, standardized $b = .76$).

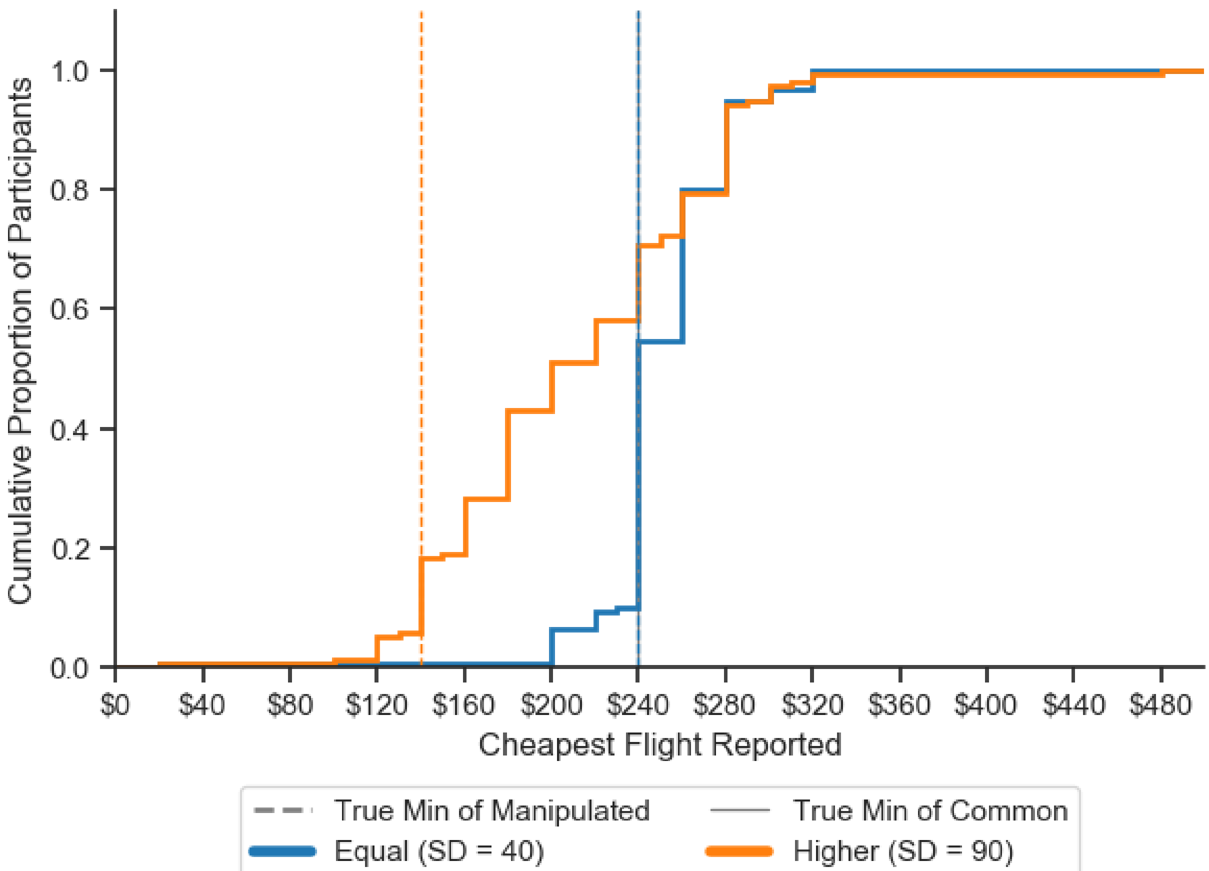


Figure 5. Cumulative distribution of cheapest flight reported in study 4.

We also find that the amount of dispersion in the manipulated category influenced participants' willingness to accept the prices offered. Figure 6 shows participants' stated likelihood to accept each offer. Two patterns emerge. First, acceptance likelihood was lower overall when dispersion in the manipulated category was higher than in the common category. This implies that dispersion spillover causes participants to view the same prices as less attractive. Second, when dispersion in the manipulated category was the same as in the common category, we observe a discontinuity in acceptance likelihood at $240 (i.e., the actual minimum price observed): Since most participants had already indicated a maximum willingness to accept a flight priced at $240, additional price reductions had a weaker impact. This implies that participants in this condition formed well-defined beliefs about the minimum price they saw. In contrast, there is no apparent discontinuity at $240 when dispersion in the manipulated category was higher, suggesting that these participants formed different—and less accurate—beliefs about the distribution of prices.

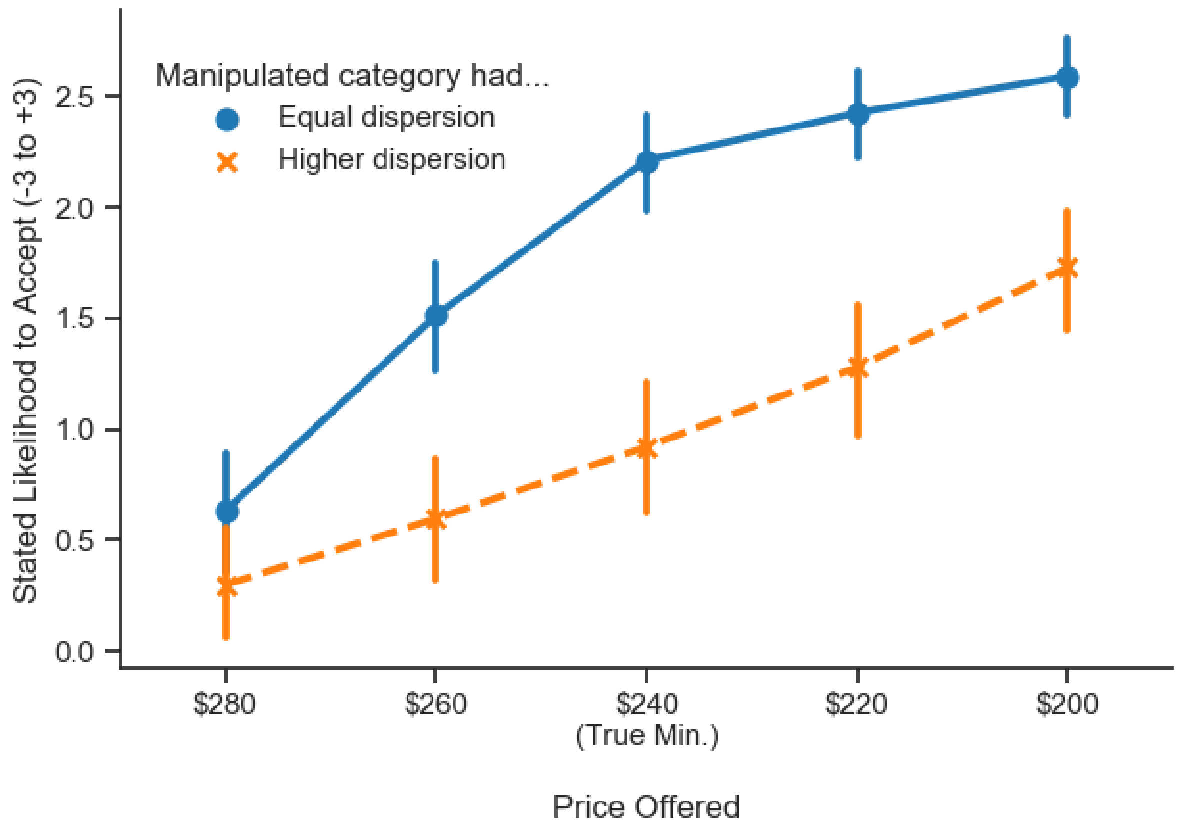


Figure 6. Average stated likelihood to accept the offers in study 4. Error bars are 95% bootstrapped confidence intervals.

To examine the statistical significance of these patterns, we analyzed participants' stated likelihood to accept the five offers with a mixed linear model that included as predictors a continuous variable coding the price of the offer ($280 vs. $260 vs. $240 vs. $220 vs. $200), a contrast variable indicating the experimental condition (equal vs. higher dispersion in the manipulated category), a contrast variable indicating whether the price of the offer was strictly below versus equal/higher than the minimum price presented in the category, all two and three-way interactions between those variables, and a random effect for each participant.

This analysis reveals several patterns. First, there is a main effect of the experimental condition. Across the five deals, participants said they were significantly more likely to accept offers when dispersion in the manipulated category was the same as in the common category ($z = -5.67$, $p < .001$, standardized $b = .52$). Second, this analysis also reveals a significant three-way interaction, confirming the discontinuity that we described above ($b = .76$, $z = 5.02$, $p < .001$, standardized $b = .43$). When dispersion in the manipulated category was the same as in the common category, the effect of a $20 price decrease on acceptance likelihood was stronger for offers above the minimum price ($z = 16.42$, $p < .001$, standardized $b = .45$) than for offers below the minimum price ($z = 1.73$, $p = .083$; two-way interaction: $z = -5.79$, $p < .001$, standardized $b = .36$). On the other hand, when dispersion in the manipulated category was higher, the effect of a $20 price decrease on acceptance likelihood was similar for offers above ($z = 6.49$, $p < .001$, standardized $b = .18$) versus below the minimum price ($z = 4.69$, $p < .001$, standardized $b = .26$; interaction: $z = 1.30$, $p = .194$).

In sum, study 4 shows that dispersion spillover has an impact on judgments of price attractiveness and intention to search.

**Study 5: Search for Lower Prices**

Study 5 demonstrates downstream consequences of dispersion spillover in a more naturalistic and incentive-compatible context. In this study, participants first received daily price notifications from a travel agent, and then booked a flight for an upcoming business trip. When booking the flight, participants saw the available price for the flight on six consecutive “days,” and each day they had to decide whether to book the flight at the price currently offered, or to wait until the next day in the hope that a better price will be offered. This sequence of decisions (akin to the "secretary problem," Bearden and Murphy 2007) is typical of circumstances in which consumers have to weigh a current offer against future expected offers (e.g., signing a lease on an apartment, booking a flight ticket or a hotel room, accepting a job offer…).

The study was incentive compatible: The more participants spent on the flight, the less they earned as a bonus. As in real life, participants tried to book the flight at the right moment to get the cheapest possible price. We predict that dispersion spillover will increase the likelihood that participants forego objectively attractive offers. We pre-registered our hypothesis, target sample size, detailed analysis plan, and expected pattern of results on as AsPredicted (Simonsohn, Simmons, and Nelson 2015). The pre-registration is available on the Open Science Framework (OSF) repository of the paper.

***Method***

Following our pre-registration, we posted 500 HITs worth 40 cents on MTurk. We obtained data from 503 respondents and excluded data from 21 participants who reported seeing a minimum price that was strictly higher than the true median price ($320), leaving a final sample size of 482. We did not record any demographic or psychographic variables.

Participants imagined they often travel for work to Florida and Colorado, and that they have subscribed to daily notifications from a travel agent indicating the best prices currently available for round-trip flights to both destinations. Participants first reviewed the 26 price notifications they had received in the past month, thus observing 26 prices for flights to both destinations (see Figure 7). This presentation format was meant to imitate "push notifications" that websites send to consumers (e.g., "SkyScanner," "Kayak," or "AirfareWatchDog").

Participants took as much time as they wanted to review each notification before dismissing it, after which a new notification appeared.

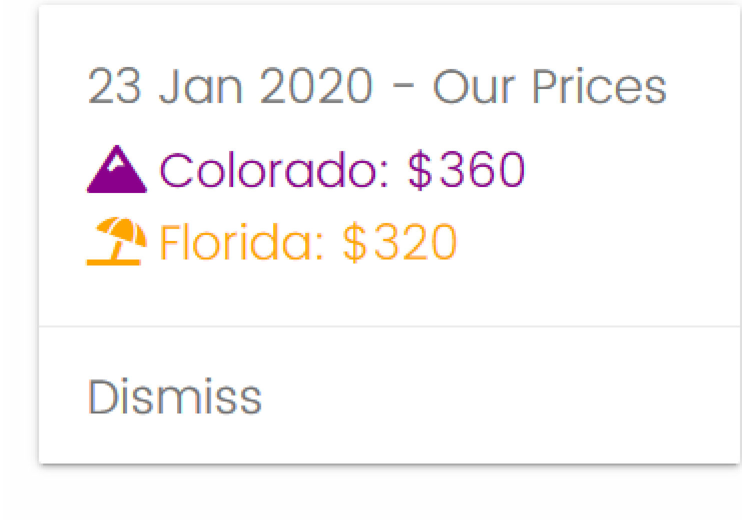


Figure 7: A daily notification shown to participants in study 5.

As in study 4, we manipulated the dispersion of flight prices that appeared on the notifications. One destination was the common category (with prices ranging between $240 and $400; SD = 40), the other was the manipulated category and had either the same price dispersion or greater (with prices ranging between $140 and $500; SD = 96). We instructed participants to pay close attention to the prices because they would have to book a flight later.

After reviewing all 26 notifications, participants learned about an unexpected business trip to Florida (or Colorado, depending on which destination was the common category) seven days from now. We told participants that they would have to book a flight with their usual travel agent. To do so, they had a travel allowance of $500, and any money remaining from this allowance could be used to enjoy meals and drinks at their destination. To make decisions incentive-compatible, we informed participants they would earn a bonus of 10 cents for every $50 they had left after booking the flight.

On each of the six "days" before departure, participants received an offer from the travel agent describing the price currently available. Each day, participants decided between booking the flight at the offered price, or delaying their reservation until the next day in the hope that a cheaper price will be offered. If a participant had not booked a flight yet the day before departure, they had to accept the price offered on that day. We further clarified that the tickets were non-refundable, that the prices offered by the travel agent would be comparable to those they had seen so far, and that the prices would not become systematically cheaper or more expensive over time.

Unbeknownst to participants, the sequence of prices offered was determined in advance and identical for all participants. The travel agent would first offer a price of $340, then a price of $320, then $260, $380, $300 and finally a price of $320 on the day before departure. Given the prices presented in the learning phase, the price of $260 offered on the third day is a clear dominant option: The expected value of rejecting this offer is negative, and there is only 11% chance of finding a lower price.

After participants purchased a flight, we finally asked them to report the minimum price they remembered seeing for the destination, and to estimate the average price of flights to this destination.

### ***Results***

As in study 4, we find that participants' beliefs about the cheapest flight they could find were affected by the amount of price dispersion encountered in the manipulated category. Figure 8 shows the cumulative distribution of minimum prices reported by participants. When price dispersion in the manipulated category was higher than in the common category (vs. equal), 43%

(vs. 8%) of participants reported a minimum price that was below the actual minimum. Following our pre-registration, we analyzed the reported minimum value using an OLS regression including as predictors a dummy for the experimental condition (equal vs. high dispersion in the manipulated category), a contrast for the counterbalancing factor (Colorado vs. Florida), and the interaction of these variables. As predicted, only the experimental condition had a significant effect ($t(481) = -6.84$, $p < .001$, standardized $b = -.60$). When dispersion in the manipulated category was higher than in the common category, the minimum price reported by participants was on average $24.85 lower.

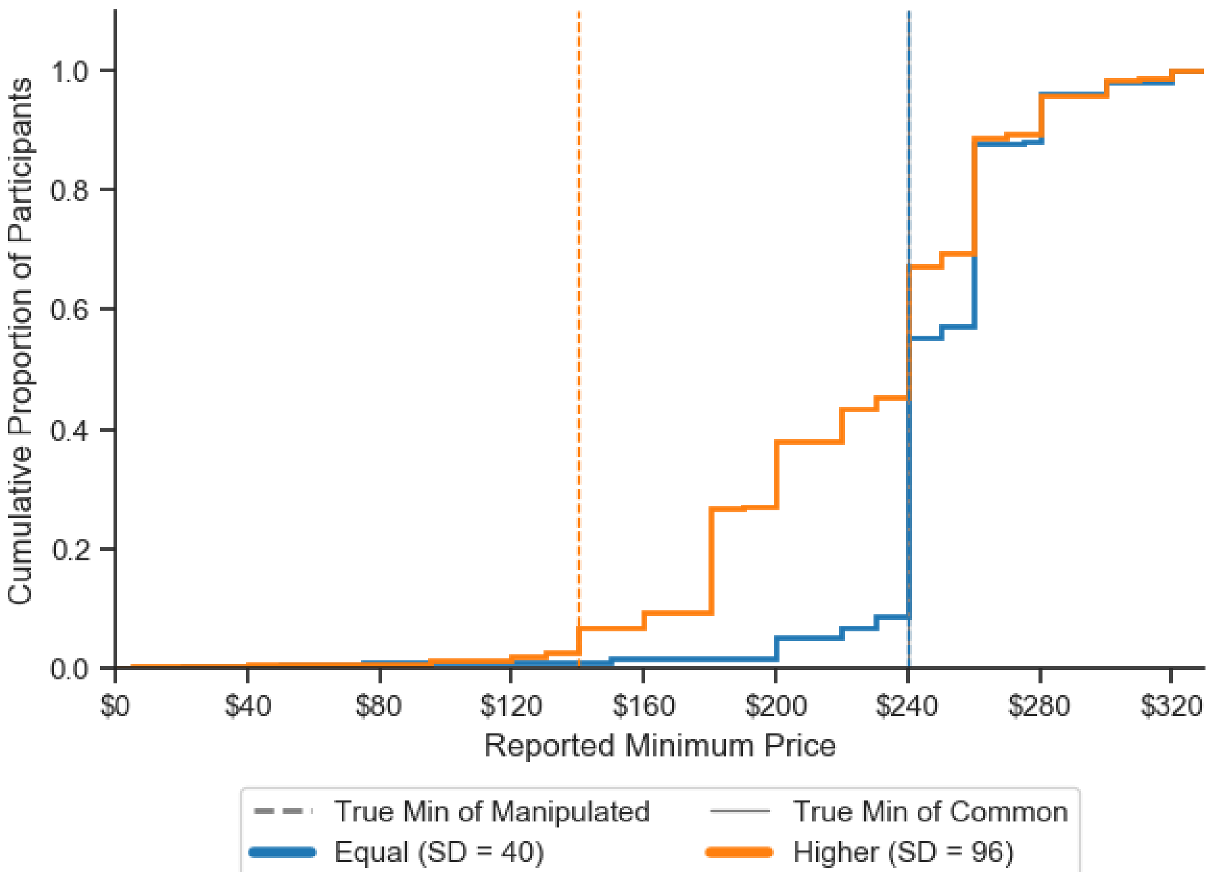


Figure 8. Cumulative distribution of minimum price reported in study 5.

We also find that the amount of price dispersion in the manipulated category had an influence on participants' search behavior. Figure 9 shows the proportion of participants accepting each of the six offers made by the travel agent. When price dispersion in the manipulated category was the same as in the common category, only 8% of participants continued searching after seeing the third offer of $260. In contrast, 21% of participants continued searching when price dispersion in the manipulated category was higher. Following our preregistration, we analyzed the duration of search using an ordered logistic model. This model represents search depth in a latent utility space divided in 5 thresholds (delineating the preference between reviewing K offers and K+1 offers), and estimates the impact of our predictors (a dummy coding the experimental condition, a contrast for the counterbalancing factor, and the interaction of those terms) in this latent utility space. As predicted, the subjective utility of search was significantly higher when dispersion in the manipulated category was high ($b$ = .47; $t(481) = 2.41$, $p$ = .016, odds ratio = 1.60).

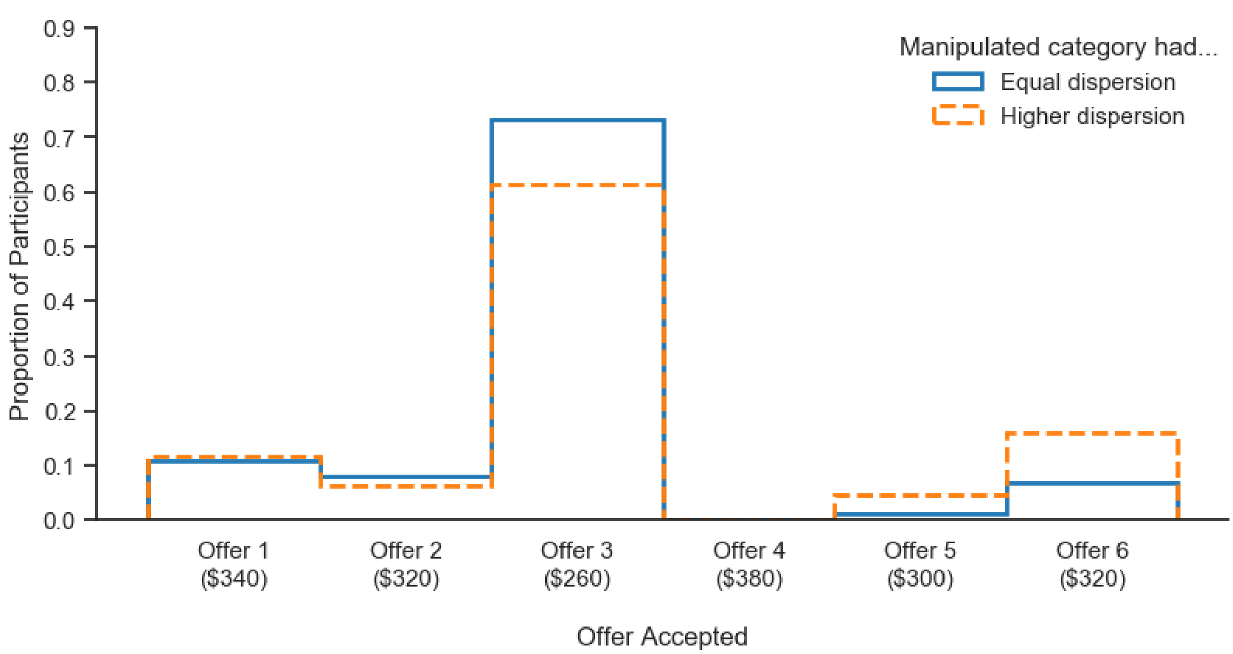

Figure 9. Proportion of participants accepting each flight price offered in study 5

In sum, study 5 demonstrates the downstream effects of dispersion spillover in an incentive-compatible setting. We found that dispersion spillover can lead consumers to forego attractive purchase opportunities.

## Study 6: Search for Higher Compensation

Many consumers today supplement their income by performing tasks in exchange for a compensation. They may drive for Uber, deliver food for Deliveroo, or complete Human Intelligence Tasks (HITs) on Amazon's Mechanical Turk platform. In the "gig economy," consumers frequently decide between accepting a current task for a given compensation (e.g., driving a passenger to the airport for $60) or rejecting it, hoping that a more attractive task will be offered. These decisions should be informed by dispersion knowledge, and thus we expect dispersion spillover to have an effect.

Study 6 examines how dispersion spillover influences the decisions of MTurk workers to accept or reject compensation offers. We predict that dispersion spillover will lead MTurk workers to forego attractive offers and earn less money as a result. We pre-registered our hypothesis, sample size, detailed analysis plan, and expected pattern of results on as AsPredicted (Simonsohn et al. 2015). The pre-registration is available on the Open Science Framework (OSF) repository.

### *Method*

Following our pre-registration, we posted 500 HITs worth 20 cents on MTurk. We obtained data from 502 respondents and excluded data from 12 participants who reported seeing a maximum bonus that was strictly lower than the true median bonus (32 cents), leaving a final sample size of 490.[3] We did not record any demographic or psychographic variables.

We told MTurk workers that we often give bonus payments to participants in our studies, and that the distribution of the bonuses we give depends on the type of study. We then presented participants with a sequence of 50 “bonuses that we have paid in the past.” Half of the bonuses were for “Red HITs” and the other half were for “Blue HITs.” We presented each bonus for 1.2 seconds. To facilitate learning, we color-coded the bonuses, and paired “Red HITs” with a circle and “Blue HITs” with a 12-pointed star.

Red and blue HITs had the same average bonus (32 cents), but we manipulated the dispersion of bonuses between participants. The common category had the same amount of dispersion across conditions (with bonuses ranging between 24 cents and 40 cents; SD = 4). The manipulated category either had the same amount of dispersion as the common category or a higher amount of dispersion (with bonuses ranging between 14 cents and 50 cents; SD = 9.6). As in previous studies, we counterbalanced the labels (i.e., color of HIT) across distributions. We did not observe a significant effect of this counterbalancing factor and do not elaborate on it further.

After viewing the bonus payments for both types of HITs, we revealed to participants the type of HIT they were currently working on (i.e., red or blue HIT, whichever was the common category). We told participants they would have an opportunity to earn a bonus and presented

[3] 48 participants entered a bonus amount that was smaller than 1. We assume that those participants entered the bonus amount in dollars rather than in cents (i.e., a value of 0.42 for a bonus of 42 cents), and recoded their responses accordingly. Excluding those participants instead leaves our key result unchanged ($p = .003$).

them with five closed boxes, numbered from 1 to 5. We explained that each box contained a possible bonus payment, drawn from the distribution of bonuses of the HIT type they were currently working on. The bonus in each box was the same for all participants: 30 cents in box 1, 38 cents in box 2, 26 cents in box 3, 20 cents in box 4, and 32 cents in box 5. After opening each box, participants could either decide to keep the bonus in the box, in which case the task would end and they would receive the bonus in the box, or decide to discard the bonus and open another box. Participants who opened all boxes received the bonus in the final box.

At the end of the study, we finally asked participants to enter the highest bonus they remembered seeing in the common category.

***Results***

We find that participants' beliefs about the largest bonus they could earn was affected by the dispersion encountered in the manipulated category. Figure 10 shows the cumulative distribution of maximum bonus payments participants reported seeing in the common category. When the manipulated category had higher dispersion, 52% of participants reported a maximum bonus payment for the common category that was higher than the actual maximum of 40 cents. Only 14% of participants did this if the manipulated category had the same dispersion.

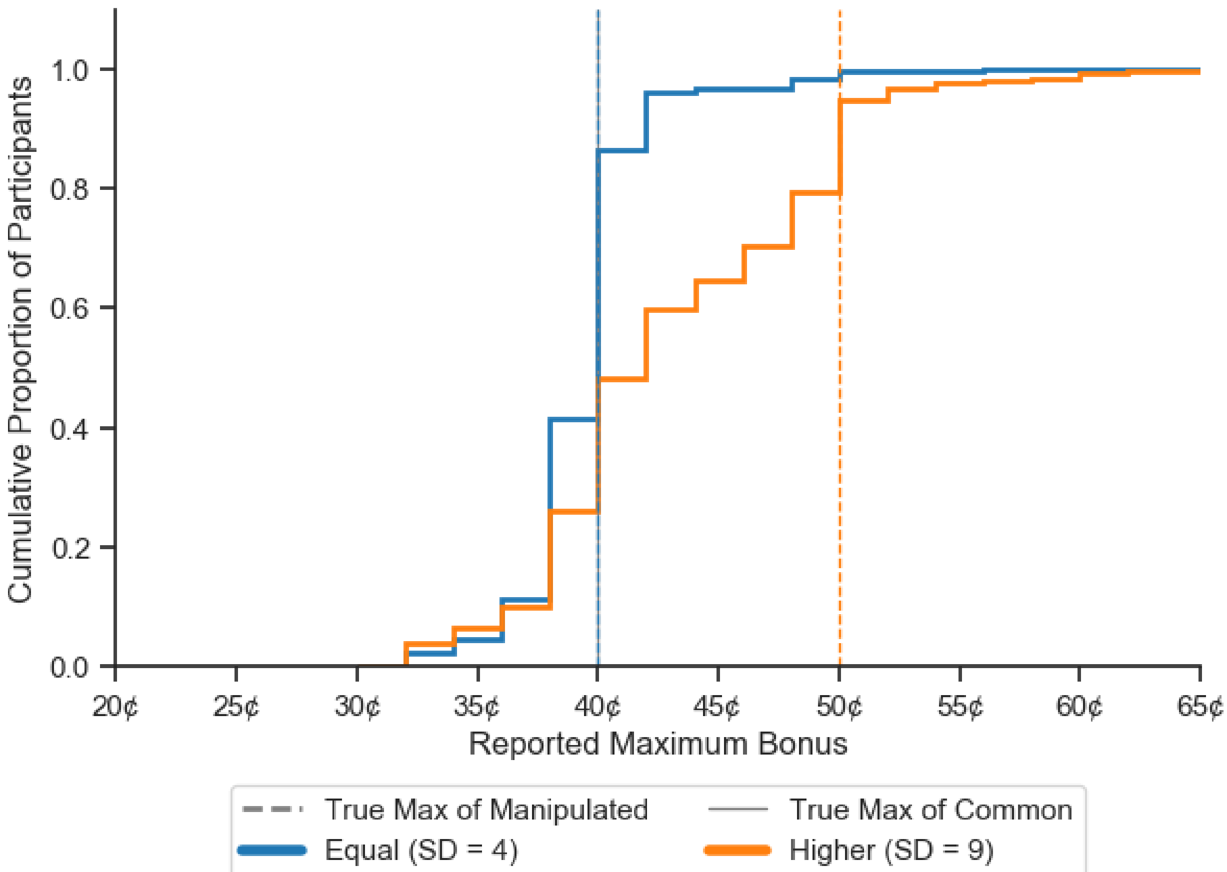


Figure 10. Cumulative distribution of reported maximum bonus in study 6

Following our pre-registration, we analyzed the reported maximum value using an OLS regression including as predictors a dummy for the experimental condition (higher vs. equal dispersion in the manipulated category), a contrast for the counterbalancing factor (red vs. blue HITs), and the interaction between these variables. As predicted, only the experimental condition had an impact. Participants reported seeing a maximum bonus that was on average 3.69 cents higher when dispersion in the manipulated category was higher versus equal ($t(486) = 8.10$, $p < .001$, standardized $b = .28$).

We also find a similar effect on participants' search behavior. Figure 11 shows the proportion of participants accepting the bonus payment in each box. Given the bonuses presented in the learning phase, the bonus of 38 cents in box 2 was very attractive: There is only an 11% chance of finding a larger bonus in one of the next three boxes. When dispersion in the manipulated category was the same as in the common category, 61% of participants accepted the bonus payment in the second box. In contrast, when dispersion in the manipulated category was higher, only 39% of participants accepted the bonus payment in the second box. More participants continued their search, ultimately receiving one of the (smaller) bonus payments contained in the following boxes.

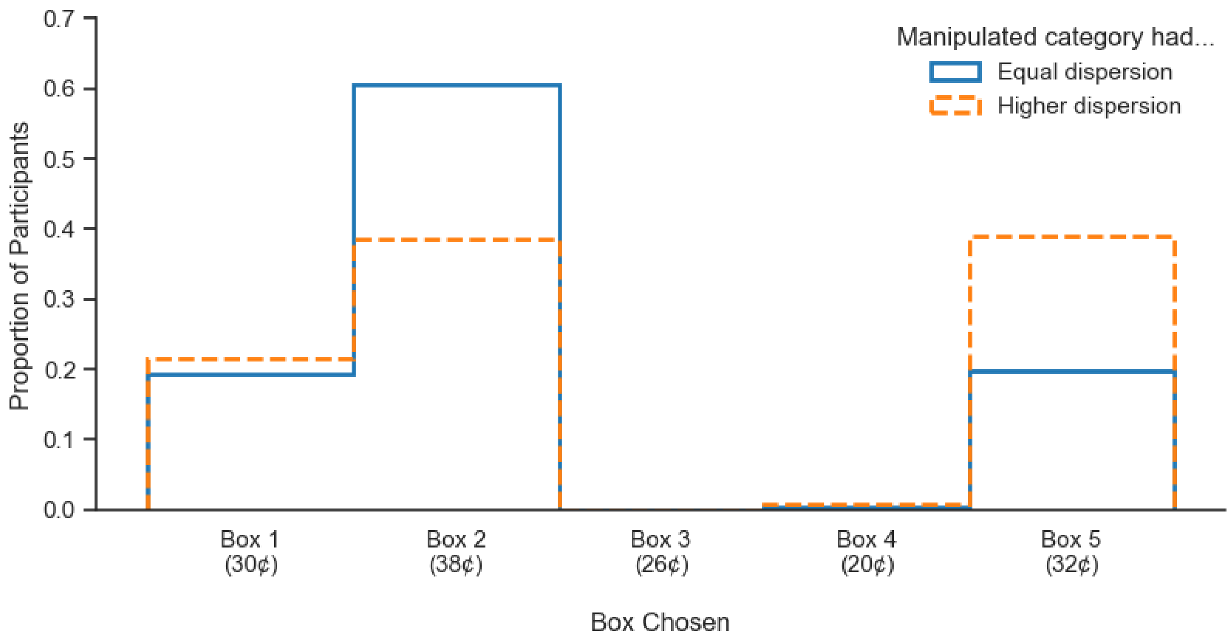


Figure 11. Proportion of participants choosing each box in study 6

Following our pre-registration, we analyzed the number of boxes opened using an ordered logistic model, as in the previous study. As predicted, the subjective utility of search was significantly higher when dispersion in the manipulated category was higher (*b* = .52; *t*(486) =

2.99, *p* = .002, odds ratio = 1.68), which led participants to earn a smaller bonus (*t*(486) = -1.42, *p* <.001, standardized *b* = -.39).

In sum, study 6 demonstrates the implications of dispersion spillover for consumer decision-making in the gig economy. We found that dispersion spillover can lead MTurk workers to forego an attractive compensation after working on a task.

## Study 7: Bidding in an Auction

Study 7 examines whether dispersion spillover influences how consumers bid in an auction. In a typical "blind" auction (also called first-price sealed auction), the optimal strategy—if you want the product—is to place a bid that is slightly higher than what you believe will be the highest bid amongst competing bidders (as long as this price is below your willingness to pay). Based on the results we obtained in previous studies, we expect that people's beliefs about the maximum bid in a particular auction will be influenced by the bids they have encountered in other auctions. When bids submitted for another item have more (vs. less) dispersion, we predict that people will overestimate the maximum bid, and as a consequence bid more themselves. We pre-registered our hypothesis, target sample size, detailed analysis plan, and expected pattern of results on as AsPredicted (Simonsohn et al. 2015). The pre-registration is available on the Open Science Framework (OSF) repository of the paper.

### *Method*

Following our pre-registration, we posted 500 HITs worth 30 cents on MTurk. We obtained data from 502 respondents and excluded data from participants who reported seeing a maximum bid that was strictly lower than the true median bid ($32), or made at least one mistake in a quiz that tested their understanding of the bidding procedure. This left us with a final sample size of 276.[4]

We first explained the rules of the auction to participants as follows: We present the item to the crowd, and all the attendants privately submit a bid. After all bids are submitted, we review the bids, and the highest bidder gets to buy the item at the price he or she entered. We then explained that we had auctioned two items so far, a $100 Amazon gift card and a $100 Whole Foods gift card. Participants reviewed the 26 bids for the Amazon gift card and the 26 bids for the Whole Foods gift card, one at a time, in two distinct blocks, as in study 3. Each bid appeared on screen for exactly one second. To facilitate learning, the bids for the Amazon gift card were displayed in orange with an icon of a shopping cart, and the bids for the Whole Foods gift card were displayed in green with an icon of a basket. We manipulated, between participants, the dispersion of the bids for the Whole Foods gift card such that they were more dispersed (bids between $14 and $50, SD = 8.6) or less dispersed (bid between $30 and $34, SD = 1.1) than the bids for the Amazon gift card (bids between $27 and $37, SD = 2.4). The median bid for both items was always $32.

After participants reviewed the bids for the two types of gift cards, we informed them that we had another $100 Amazon gift card to auction. We explained that 25 other people had already submitted their bids, and that these bids are similar to the ones they saw earlier for the other Amazon gift card. We then endowed each participant with $60, and gave them the choice to bid any amount from this money to try to win the Amazon gift card. We informed all respondents

[4] Our exclusion criteria turned out to be more stringent than expected. A robustness check ran on the full sample (N = 502, reported in the online appendix) also supports our key hypothesis (p = .002)

that one participant would be randomly selected, and that his or her choice will be enacted for real.

After participants submitted their bid (e.g., $40), we again explained the rules of the auction. If their bid was higher than the other 25 bids, they would pocket the $100 Amazon gift card, plus any leftover money (e.g., $20). Conversely, if any of the 25 bids is higher than their own, they would simply pocket the $60. After reading those instructions, participants could validate their bid, or go back and change it. Finally, we asked participants to report the highest bid (apart from theirs) they saw for the Amazon gift card, and we tested their understanding of the bidding procedure.

***Results***

Again, we find that the dispersion of bids for the Whole Foods gift card had an impact on people's memory for the maximum bid submitted for the Amazon gift card. Figure 12 shows the cumulative distribution of the maximum bids reported by participants. When bids for the Whole Foods gift card had a larger amount of dispersion, 27% of participants reported a maximum bid that was above the true maximum bid of $37. Only 3% of participants did so when bids for the Whole Foods gift card had a smaller amount of dispersion (27% vs. 3%: $\chi(1) = 33.74$, $p < .001$). Participants reported seeing a maximum bid for the Amazon gift card that was on average $1.84 higher when the dispersion of bids for the Whole Foods gift card was higher ($t(274) = 5.46$, $p < .001$, standardized $b = .63$).

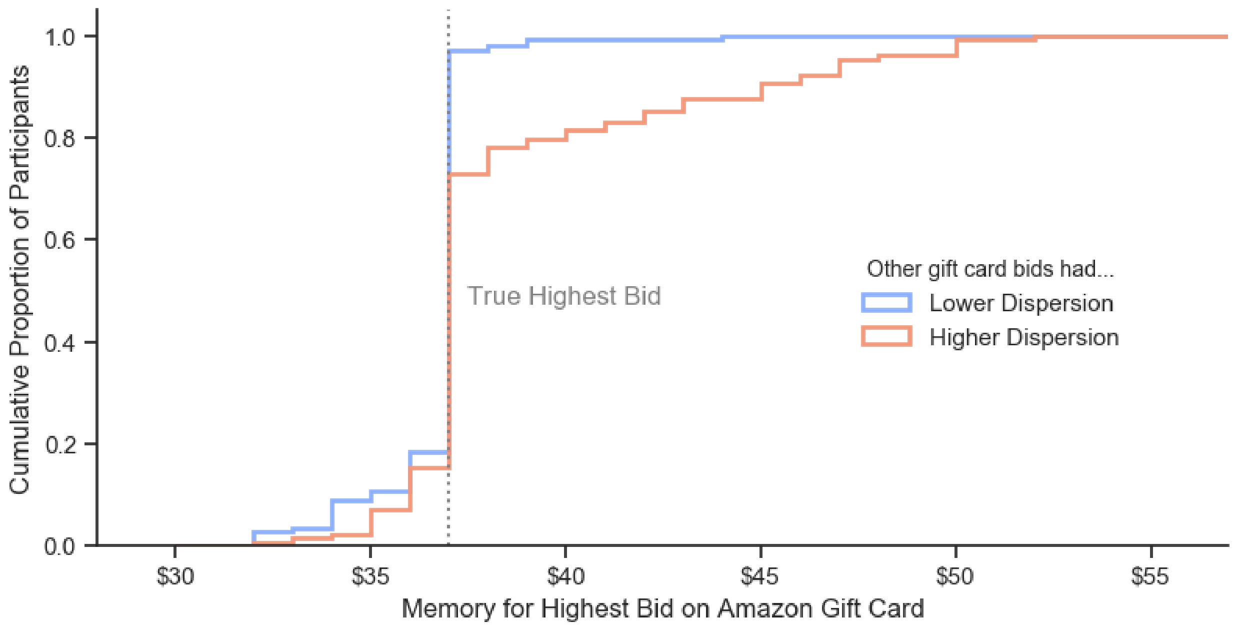


Figure 12. Cumulative distribution of reported highest bid for Amazon gift card in study 7

We observe a similar impact on the bids submitted by participants (shown in Figure 13). On average, participants bid $2.60 more when the dispersion of bids for the Whole Foods gift card was high versus low ($t(274) = 2.46$, $p = 0.014$, standardized $b = .30$).

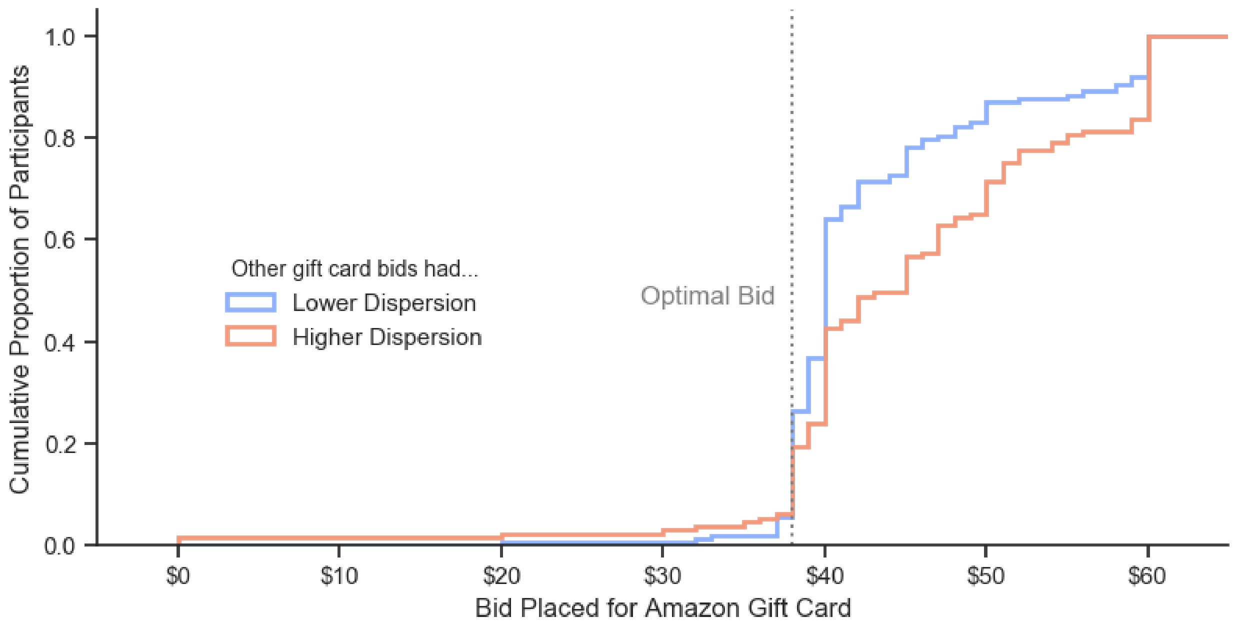


Figure 13. Cumulative distribution of bids offered in study 7

In sum, study 7 demonstrates that dispersion spillover can lead consumers to form inaccurate beliefs about how much money other people would bid, and in turn lead them to bid excessively in auctions.

***Replication with College Students***

We conducted an exact replication of this study with students enrolled at a US university and a Dutch university. We invited students to take part in the study for partial course credit and the opportunity to win a gift card. We had to change the experimental protocol described in the pre-registration in two ways. First, for the Dutch students, we substituted the gift card from Whole Foods with a gift card from Bol.com. Second, we were unable to recruit our target of 500 students, and only managed to obtain 493 responses. Applying the pre-registered exclusion criteria left us with 325 valid responses. Replicating the results we found for MTurk workers, students reported seeing a maximum bid for the Amazon gift card that was on average $1.61 higher when the dispersion of bids for the other gift card was higher ($t(323) = 5.04$, $p < .001$, standardized $b = 0.54$). Students also submitted bids for the second Amazon gift card that were on average $2.88 higher when dispersion of bids for the other gift card was higher ($t(323) = 3.11$, $p = .002$, standardized $b = 0.34$).

## PART 3: MENTAL REPRESENTATIONS OF DISPERSION

### Two Types of Mental Representations

In part 3, we explore the mental representations that might underlie the dispersion spillover we observed in previous studies. Cognitive scientists have distinguished between two types of mental representations that can support judgments: (1) concrete *exemplars* that people have seen before (e.g., Medin and Schaffer 1978), and (2) abstract summary representations, often called *prototypes* or *rules*, that people have induced from previous experience (e.g., Hampton 1979; Rosch and Mervis 1975; Smith and Medin 1981).

Our studies examine how people judge numeric distributions. In this context, exemplars are the specific values that people have encountered in each category. Prototypes, on the other hand, would correspond to higher-level representations that people have abstracted away from the specific values they have seen. They are “intuitive statistics” that summarize the distributions (Gigerenzer and Murray 2015; Lindskog, Winman, and Juslin 2013). Since people typically expect quantities to be normally distributed (Crosetto et al. 2020; Flannagan, Fried, and Holyoak 1986; Fried and Holyoak 1984), they might intuitively represent a numerical distribution using an “intuitive average” and an “intuitive dispersion” (as a normal distribution is sufficiently described by its mean and standard deviation).

It is clear that exemplars are part of distribution knowledge. We all have memories for events we have experienced. In fact, the asymmetry in the dispersion spillover that we have observed in studies 1–3 (i.e., tight distributions do not shrink the perceived dispersion of wide distributions) supports the idea that distribution knowledge is based in part on exemplars. Indeed, if a person has some memories of the extreme prices that appeared in a wide distribution, those memories are unlikely to be erased by the fact that the extreme prices did not appear in the tight distribution, and the person will make appropriate judgments about the dispersion of the wide distribution.

Identifying whether judgments are based uniquely on exemplars, or also involve intuitive statistics, is a deep theoretical issue, and one that is difficult to tackle empirically. In many contexts there is controversy about whether the two types of mental representations can be disentangled (e.g., DeLosh, Busemeyer, and McDaniel 1997; Hahn and Chater 1998; Reber 1989; Shanks and John 1994). However, in the context of dispersion spillover, we believe the two models make different predictions that can be tested empirically.

**Which Mental Representations Underlie Dispersion Spillover?**

How would prices encountered in one category change people’s perception of prices in another category? It is possible that imperfect exemplar memory plays a role. People may misattribute prices observed in one category to the other, akin to source neglect (Johnson, Hashtroudi, and Lindsay 1993). For instance, you might correctly remember seeing a price of

$15, but incorrectly remember it was for a pillow (instead of a blanket). Beyond this confusion, it is possible that when people make judgments about the distribution of prices in one category, memory traces from other categories are also activated and weighted in judgments. If the dispersion spillover we observed in previous studies is entirely driven by exemplars, we should be able to identify other spillover effects. For instance, when presented with distributions that have similar variances but different means, you would expect people's judgments to reflect an average of both distributions. That is, that people would create distributions with similar means (i.e., the perceived means of the two distributions lie closer together than the actual means) and with exaggerated variances (i.e. the perceived variance of a distribution is higher than its true variance). We test this prediction in study 8.

Pure reliance on exemplar memory makes another prediction. Exemplar-based judgments are formed by activating specific exemplars in memory. While different computational models make different assumptions about which exemplars are retrieved and how they are combined to form a judgment (e.g., Gillund and Shiffrin 1984; Hintzman 1984; Juslin, Winman, and Hansson 2007; Juslin et al. 2007; Kareev et al. 2002; Murdock 1995), an important commonality between these models is that judgments are constrained by the range of previously seen exemplars (DeLosh et al. 1997; Hahn and Chater 1998; Juslin et al. 2008). In other words, if dispersion spillover is purely based on exemplar memory, we should not find that people systematically recall seeing values outside of the range of values presented across the two categories. Observing such "phantom prices" would signal that people generate prices not only based on exemplar memory, but also based on intuitive statistics. We test this prediction in study 9.

**Study 8: Manipulating Means**

Study 8 uses a similar design to that of study 1. All participants learned about a common category and about a second, manipulated, category. Instead of manipulating the dispersion of prices in the manipulated category, as we did in study 1, we manipulated the mean price in this category to be lower, equal, or greater than the mean price in the common category.

If the dispersion spillover is uniquely driven by exemplar memory, we should observe two patterns. First, when the means of the categories are different, participants should exaggerate the dispersion of each category. For instance, if white wines are priced between $16 to $34, and red wines between $26 to $44, mistaking a $38 red wine for white wine, and a $22 white wine for a red wine, would inflate the perceived dispersion in both categories.

Second, we should observe a spillover of means, similar to the spillover of dispersions we observed in previous studies. For instance, in the previous example, mistaking a $38 red wine for a white wine would increase the mean of the white wine distribution. We note, however, that such mean spillover would be inconsistent with past research documenting people's ability to form category-specific perceptions of averages (e.g., Chong and Treisman 2005; Malmi and Samson 1983).

***Method***

We paid 150 respondents from MTurk 60 cents for their participation. Due to a technical glitch, we were unable to record data from one respondent, leaving a final sample size of 149. We did not record any demographic or psychographic variables. The learning phase was identical to study 1, except for the distribution of prices. The price dispersion was identical for the common and the manipulated category (SD = 4.5), and we manipulated the mean prices between participants. The common category had a mean price of $25, and the manipulated category had a mean price of either $15, $25, or $35 depending on condition. In the test phase, participants entered the 25 values that they remember seeing for each of the two distributions on two separate distribution builders.

### *Results*

As in study 2, each participant created two distributions in the test phase, one for the manipulated category, and one for the common category. Figure 14 shows the average means (top panel) and range (bottom panel) of these price distributions. The top panel shows that the averages of the distributions created for the common category were close to the actual mean of $25, with little or no influence of the manipulated category. Moreover, the averages of the distributions created for the manipulated category closely approximated the actual means of $15, $25, and $35. The bottom panel shows that participants reported a similar range for both categories, regardless of whether the two categories had the same mean (middle bar) or different means (other two bars). Thus, people's perceptions of the mean and dispersion of prices in the common category do not appear to be significantly affected by the mean price in the manipulated category.

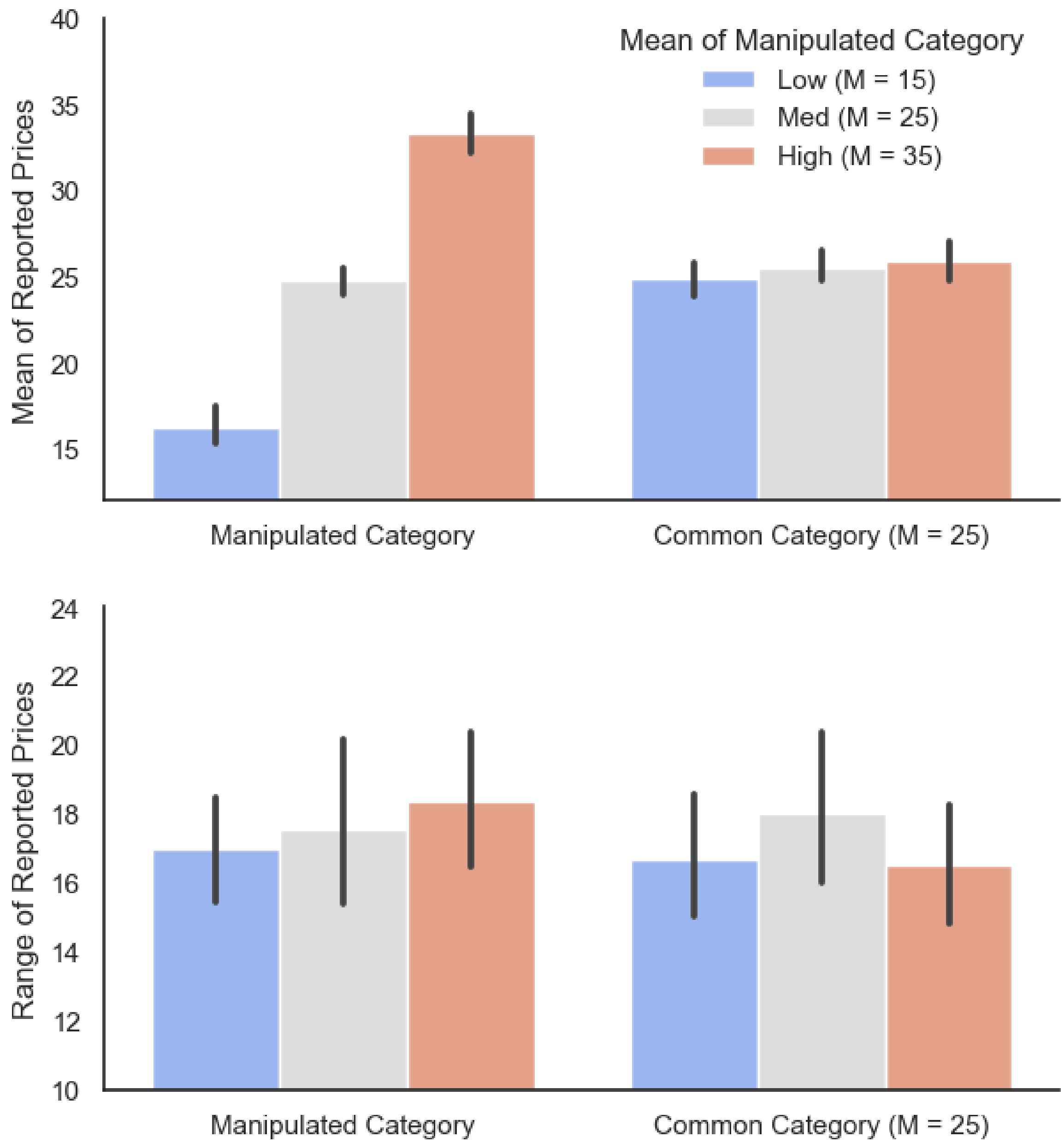


Figure 14. Average means (top panel) and ranges (bottom panel) of distributions created by participants in study 8. Error bars are 95% bootstrapped confidence intervals.

To examine statistical significance, we estimated a mixed linear model with category type (common vs. manipulated) as a within-participant factor, the actual mean of the manipulated category (low vs. medium vs. high) as a between-participant factor, and the interaction between these two factors.

We first analyzed the means of the reported distributions. For the manipulated category, participants appropriately created distributions with higher means if the actual mean price in the manipulated category was high versus medium ($M_{high}$ = 33.36 vs. $M_{med}$ = 24.73: z = 11.90, $p$ < .001, standardized $b$ = 1.40), and medium versus low ($M_{med}$ = 24.73 vs. $M_{low}$ = 16.27: z = 11.44, $p$ < .001, standardized $b$ = 1.38). For the common category, participants appropriately created distributions with means that were not statistically significantly different from each other ($M_{high}$ = 25.95 vs. $M_{med}$ = 25.54 vs. $M_{low}$ = 24.86, $p$ = .134 for largest pairwise difference).[5] We then analyzed the ranges of the reported distributions. We find that participants created distributions with comparable ranges across conditions, both for the manipulated category ($M_{high}$ = 18.38 vs. $M_{med}$ = 17.55 vs. $M_{low}$ = 17.00, $p$ > .332 for all pairwise differences) and for the common category ($M_{high}$ = 16.54 vs. $M_{med}$ = 18.00 vs. $M_{low}$ = 16.67, $p$ > .304 for all pairwise differences).

We conclude that the dispersion spillover is unlikely to be uniquely driven by a misattribution of prices from one category to the other, or the mental activation of prices across the two categories. Exemplar confusion or activation between categories should have produced a mean spillover, but participants in study 8 created distributions with means that closely approximated the actual means for both the manipulated and the common category. This finding conceptually replicates previous investigations of how people learn central tendencies in multi-category environments (e.g., Chong and Treisman 2005; Malmi and Samson 1983), bolstering the validity of our learning paradigm and dependent measures. Moreover, exemplar confusion or activation between categories should have led people to create wider distributions when the

---

[5] In the online appendix, we present multiple results showing that this pattern is more consistent with the null than with other alternatives.

actual means of two categories are different, but the ranges of the distribution participants created were similar across conditions.

**Study 9: Phantom Prices**

Study 9 provides more direct evidence that dispersion spillover is driven in part by the formation of "intuitive dispersions." Participants in study 9 learned about two categories with non-overlapping price distributions: Regardless of their dispersions, the maximum price in one category was strictly lower than the minimum price in the other category. In line with previous studies, we expect participants to perceive prices in a given category as more dispersed if the other category had a larger amount of dispersion. Because the two distributions do not overlap, such dispersion spillover would lead participants to report more *phantom prices* that they have never encountered in any of the distributions. Such increase in the number of phantom prices cannot be accounted for by a simple misattribution of prices from one category to another.

We further distinguish between two types of phantom prices, those that lie in between the two distributions (interior), and those that lie outside the range of prices presented across both distributions (exterior, see Figure 15 for an illustration). The observation that people report more interior phantom prices after learning a high (vs. low) amount of dispersion in another category could be explained by an exemplar-based model of judgment, even if the price distributions do not overlap. For instance, suppose prices of white wines range from \$8 to \$16 (as in Figure 15). From an exemplar-based perspective, an interior phantom price for white wines (e.g., \$22) would result from the accidental activation of a red wine price (e.g., \$28) and its combination with a white wine price (e.g. $\$16 \times .5 + \$28 \times .5 = \$22$). If the distribution of red wines is more dispersed (i.e. starts at $28 rather than $33), the greater proximity (i.e., similarity, Casasanto 2008; Rips 1989) of the red wines to the white wines could facilitate their activation, and increase the number of interior phantom prices that participants report.

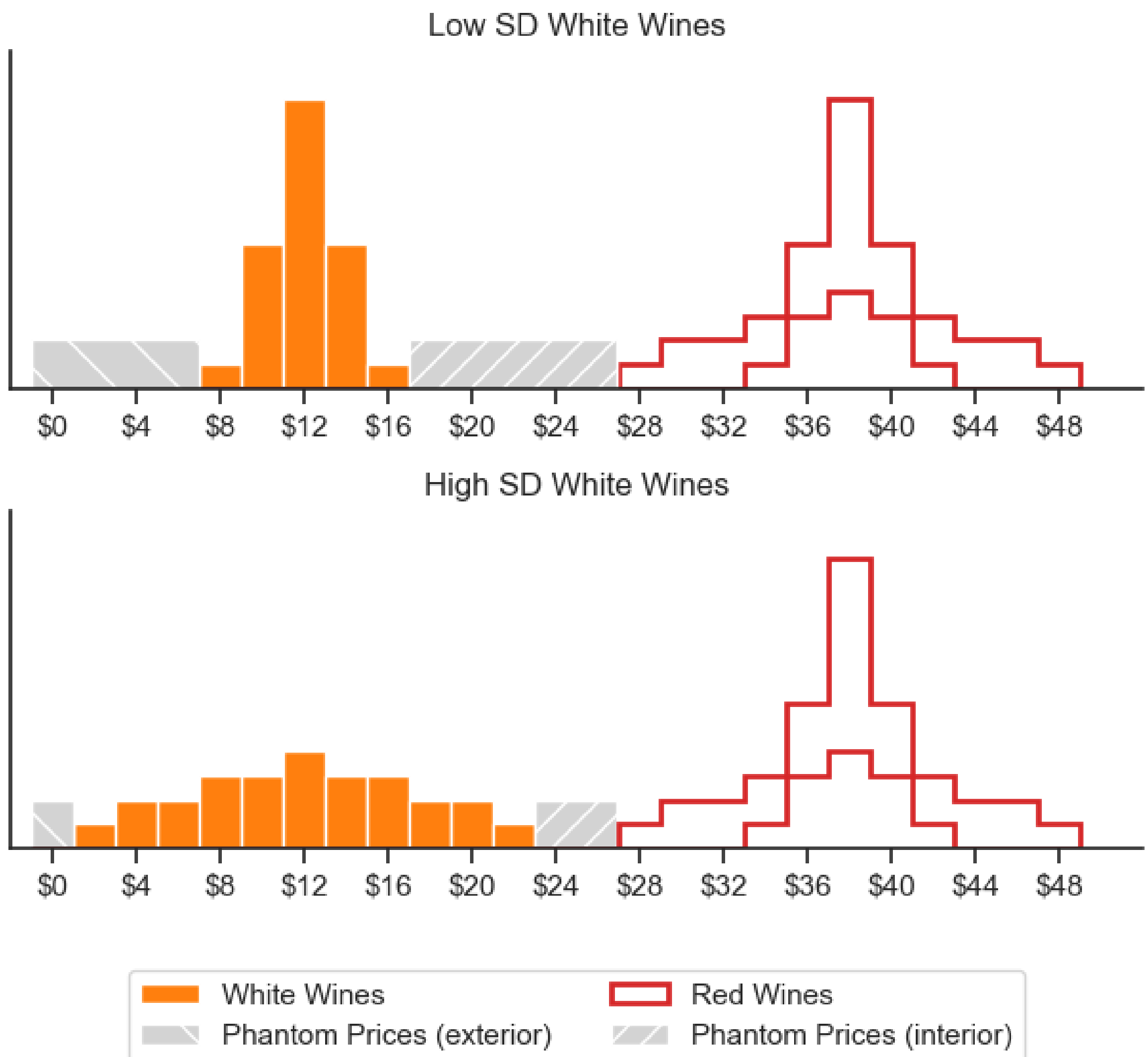


Figure 15. Visual representation of exterior and interior phantom prices for white wines. A similar logic is applied to red wines.

A greater number of *exterior* phantom prices, on the other hand, would be difficult to explain with exemplar-based models of judgment alone. Following our previous example, a white wine priced at $4 cannot be described as a linear combination of prices observed across the two wine categories (DeLosh et al. 1997; Hahn and Chater 1998; Juslin, Olsson, and Olsson 2003). If we observe that people report a greater number of exterior phantom prices when the other distribution has high (vs. low) dispersion, it would mean that the prices that people report are informed by intuitive statistics of dispersion.

Besides the non-overlapping nature of the distributions, another feature of the design warrants attention. While in previous studies we manipulated the properties of one category holding constant the other, in this study we orthogonally manipulated the price dispersion of both categories (see Figure 16). This allows us to quantify the extent to which people's price judgments for a given category are influenced by the actual price dispersion in *this* category versus the price dispersion in the *other* category.

### ***Method***

We paid 304 respondents from MTurk 60 cents for their participation. We did not record any demographic or psychographic variables. Data from three participants were not recorded because of a technical glitch, so the final sample size for analysis was 301. All participants saw 26 prices for red wines and 26 prices for white wines, in random order. The mean price of red wines and white wines was the same for all participants ($M_{\text{red}}$ = $38 vs. $M_{\text{white}}$ = $12), but we orthogonally manipulated the standard deviations of the two price distributions to be low (SD = 1.75) or high (SD = 5.26). Figure 16 illustrates the prices participants saw in the different conditions. The learning phase was identical to that of study 8, and we again used the distribution builder tool to measure participants' knowledge for the prices presented in the learning phase.

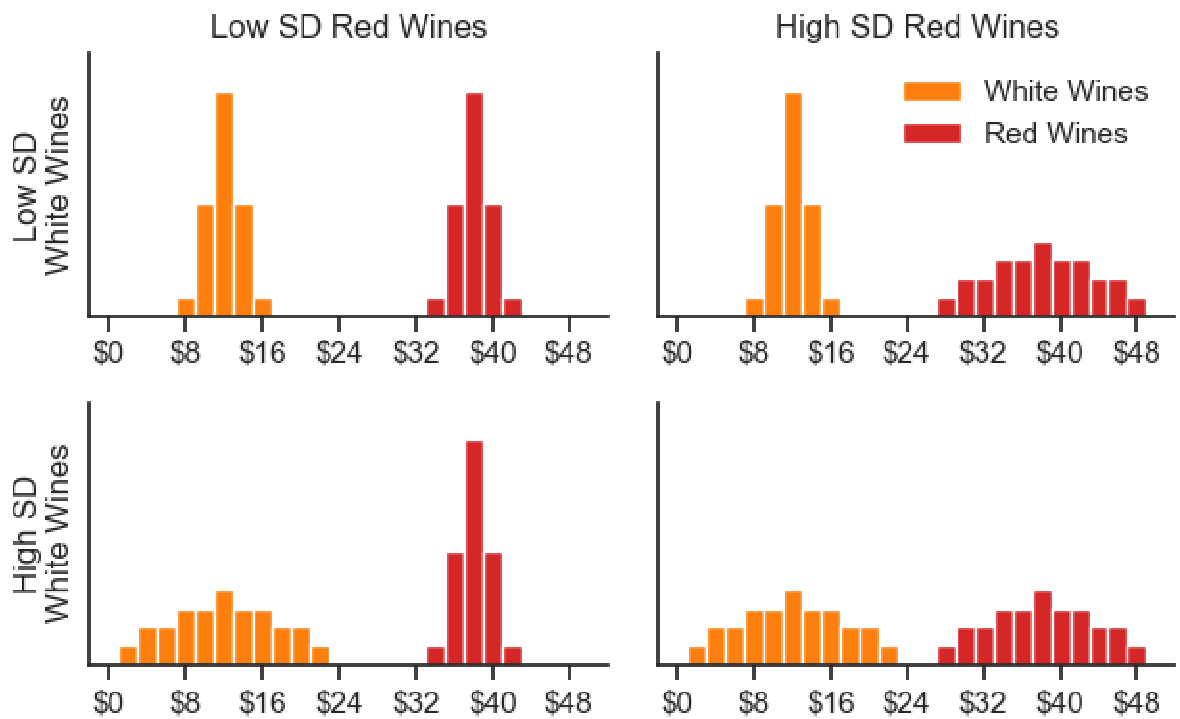


Figure 16. Distributions of prices presented to participants in study 9.

***Results***

Figure 17 visualizes the standard deviations of the distributions that participants created. It reveals two effects: First, when a category had higher price dispersion, participants appropriately created a more dispersed distribution of prices for that category. This can be seen by comparing the average height of the left two bars with the average height of the right two bars. Second, we replicate the dispersion spillover: Participants inappropriately reported a more disperse distribution of prices for a category when the other category had higher price dispersion. This can be seen by comparing the bars of different colors to each other.

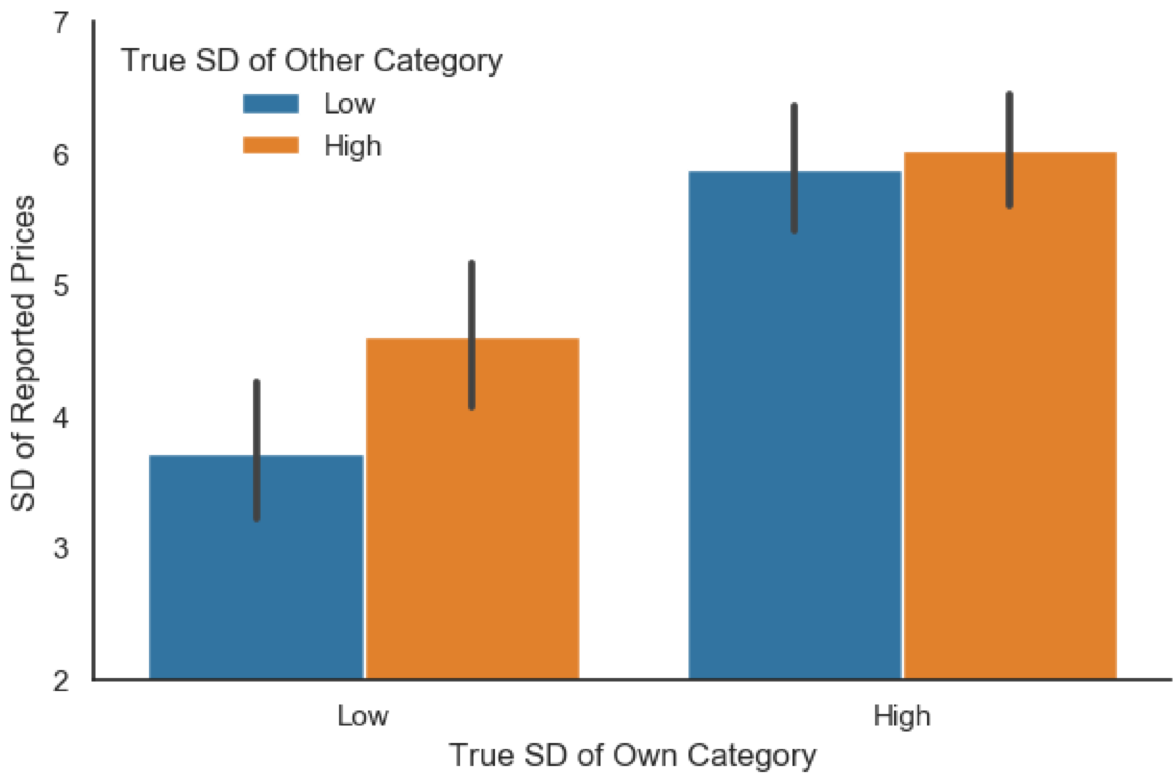


Figure 17. Average standard deviations of distributions created by participants in study 9. Error bars are 95% bootstrapped confidence intervals.

We assessed statistical significance using a mixed linear model that regressed the standard deviation of each reported distribution on a contrast variable indicating whether the *own* price dispersion of the category was actually high (vs. low), a contrast variable indicating whether the price dispersion in the *other* category was actually high (vs. low), the two-way interaction of the contrast variables, and a random intercept for each participant.

This analysis revealed a significant effect of a category's own dispersion. The average standard deviation of the reported distribution was 1.79 larger for categories with a high (vs. low) amount of actual price dispersion ($z$ = 6.94, $p$ < .001, standardized $b$ = 0.54). The analysis also revealed a significant effect of the other dispersion, which was about a third of the size of the coefficient for the own dispersion. The average standard deviation of the reported distribution for a category was 0.52 larger when the *other* category presented in the learning phase had a high (vs. low) price dispersion ($z = 2.00$, $p = .046$, standardized $b = 0.16$). We again find that the prices participants generated for a given category were a function of both the actual price dispersion in that category and also the price dispersion in a category that was learned concurrently. The interaction effect between a category's own dispersion and the other category's dispersion was not significant ($p > .22$).

We also counted how many phantom prices participants reported in their distributions. If phantom prices were uniquely driven by inattention or random responses, we should observe that participants reported a similar number of them regardless of whether the other dispersion had low dispersion or high dispersion. Instead, we observe that people reported 1.98 additional phantom prices when the other distribution had high (vs. low) price dispersion (z = 4.345, $p < .001$, standardized $b = .48$).

As mentioned earlier, a stricter test of the influence of intuitive statistics of dispersion would only consider the "exterior" phantom prices (i.e., values that were strictly lower or higher than all the prices presented across both distributions). We found that people reported 0.44 additional "exterior" phantom prices when the other distribution had high (vs. low) price dispersion (z = 2.183, $p = .029$, standardized $b = .25$).

In sum, study 9 demonstrates dispersion spillover when two distributions do not overlap. Participants reported more exterior phantom prices when price dispersion in another category was higher. This suggests that consumers' judgments of distributions are not only based on exemplar memories, but they are also informed by intuitive statistics of dispersion that are not independent across categories.

**Study 10: Category Similarity**

The previous studies suggest that the dispersion spillover can be explained by the formation and application of intuitive dispersions that are not entirely category-specific. This spillover might reflect a form of ecological rationality (Gigerenzer and Selten 2002). It is reasonable to assume that distributions presented in the same environment have statistical similarities, and thus pooling dispersion information across multiple categories may be an adaptive strategy.

Study 10 examines the role of category similarity. The study has a very similar design to that of study 9. We orthogonally manipulated price dispersions in two categories, but the distributions from both categories now overlap partially (see Figure 18). We also manipulated category similarity between participants. This allows us to compare the degree to which people assimilate the price dispersion of two product categories when those categories are more similar (i.e., red wines and white wines) versus dissimilar (i.e., red wines and smartphone cases).

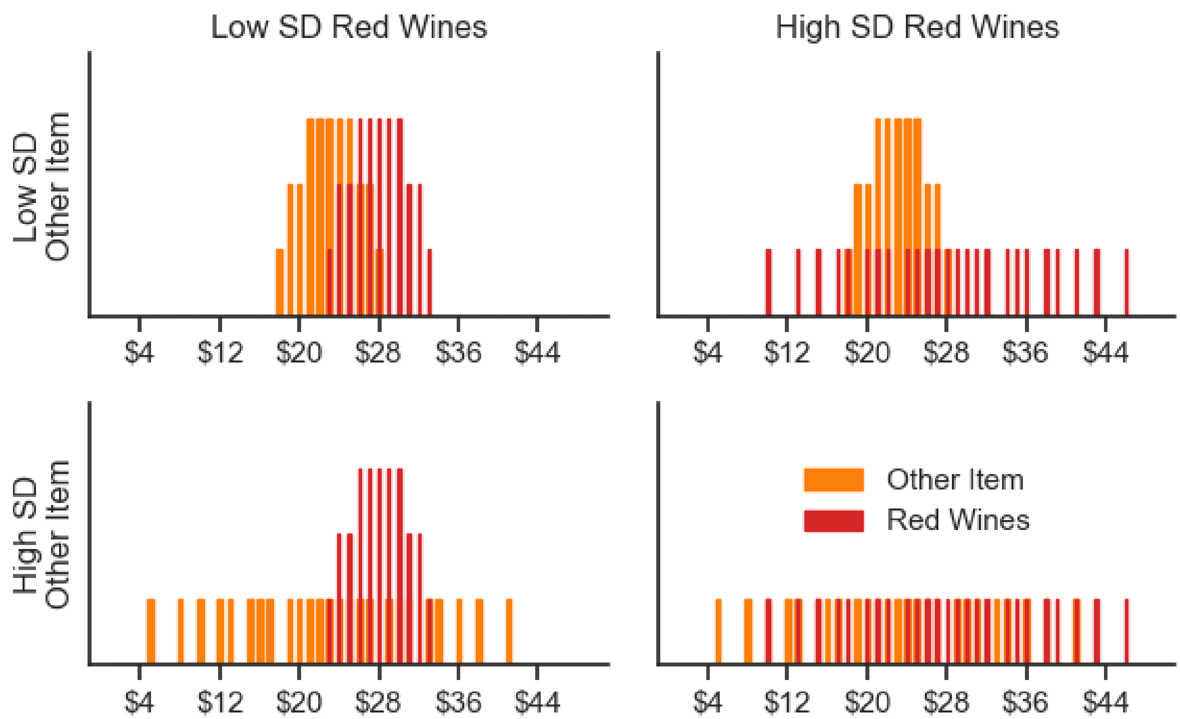


Figure 18. Distributions of prices presented to participants in study 10.

***Method***

We paid 900 respondents from MTurk 60 cents for their participation. This increase in per-cell sample size compared to study 9 was based on recommendations for testing attenuation effects (Simonsohn 2014). We did not record any demographic or psychographic variables. Data from 23 participants were not recorded because of a technical glitch, so the final sample size for analysis was 877. We randomly assigned participants to one of eight cells of a 2 (price dispersion of red wines: low vs. high) × 2 (price dispersion of other category: low vs. high) × 2 (category labels: red wines and white wines vs. red wines and smartphone cases) factorial design. The first two factors were identical to study 9, and the third factor manipulated the labels and pictures paired with the prices so that they described similar categories (red and white wines) or dissimilar categories (red wines and smartphone cases). The test phase for study 10 was identical to that of study 9.

***Results***

To analyze the data, we used the same analytical strategy as in study 9, adding category similarity as a factor in the mixed linear model. The model thus regressed the standard deviation of each reported price distribution on a contrast variable indicating whether the own price dispersion of the category was high (vs. low), a contrast variable indicating whether the price dispersion of the other category in the learning phase was high (vs. low), a contrast variable indicating whether the product categories were dissimilar (vs. similar), all two-way and three-way interactions between those variables, and a random intercept for each participant.

As in study 9, the analysis revealed not only a significant effect of the category's own dispersion ($b$ = 3.45, z = 35.68, $p$ < .001, standardized $b$ = 1.26), but also of the dispersion of the other category ($b$ = 1.25, z = 12.75, $p$ < .001, standardized $b$ = .45), which replicates dispersion spillover. In this study, the increase in statistical power allowed us to detect a two-way interaction between a category's own dispersion and the other category's dispersion ($b$ = -0.620, z = -2.70, $p$ = .007, standardized $b$ =-.22). This result is consistent with the asymmetry we observed in previous studies, suggesting that the perceived dispersion of a distribution is more affected when the other distribution has more (vs. less) dispersion.

The effects of a category's own dispersion and the other category's dispersion were qualified by interactions with category similarity (own dispersion × similarity: $b$ = -.47, z = 2.43, $p$ = .016, standardized $b$ = -.17; other dispersion × similarity: $b$ = .39, z = -1.93, $p$ = .054, standardized $b$ = .14). Together, these two-way interactions support the predicted moderation:

When similarity was low (vs. high), a given category's actual price dispersion had a stronger influence on perceived price dispersion and the *other* category's actual price dispersion had a smaller but significant ($z = 7.73$, $p < .001$, standardized $b = .38$) influence on perceived price dispersion.

In sum, study 10 shows that category similarity moderates dispersion spillover. This result suggests there may be some ecological rationality to dispersion spillover. It might reflect a trade-off between accuracy and effort (Payne, Bettman, and Johnson 1997).

## GENERAL DISCUSSION

### Summary

Ten studies examined how price dispersion in one category influences consumer judgments and decisions in another category. Across several incentive-compatible experiments, we demonstrated that more-dispersed prices one category decreases the perceived attractiveness of prices in another category, decreases the likelihood that consumers search for (and find) better prices, and increases how much they bid in auctions.

While participants in our studies formed relatively accurate perceptions of dispersions, and they often accurately reported the minimum and maximum prices they saw, participants' knowledge of dispersion was not entirely category-specific. This dispersion spillover emerged regardless of whether prices were presented individually or in pairs, categories were presented simultaneously or sequentially, distributions overlapped completely, partially, or not at all. It emerged regardless of how we labelled the product categories (wines, pillows, blankets, flights, gift cards, bonus amounts), although it was more pronounced when categories were similar. Across studies, we used various measures of dispersion (minimum and maximum, standard deviation, and range) and we used distribution builders to elicit people's perceptions of entire distributions. We observed the dispersion spillover regardless of which dispersion measure or elicitation method we used.

The dispersion spillover cannot be simply explained by the misattribution or confusion of prices across categories (i.e., remembering an observed price, but associating it with the wrong category). If this were the case, we should have found that learning about a distribution with a higher mean increased the perceived mean and the perceived dispersion of another distribution with a lower mean, but we could not find such evidence. We did find that people reported *phantom prices* that were lower or higher than the minimum and maximum prices encountered across both categories. This suggests that people form intuitive statistics of dispersion to summarize distributions that are not entirely category-specific.

**Intuitive Means and Dispersions in Multi-Category Environments**

While our results suggest that intuitive statistics underlie dispersion spillover, it remains an open question *why* these abstract summary representations are more category-specific for central tendency than for dispersion.

We suggested in study 10 that dispersion spillover may reflect an adaptive feature of cognition (Gigerenzer and Murray 2015; Gigerenzer and Selten 2002). Tracking variance online is difficult, so it may be functional to aggregate dispersion information across distributions, both to conserve working memory and reduce estimation bias. It is also noteworthy that, for small samples, the sample variance is a noisy and biased estimate of the population variance. If related categories are expected to have a similar amount of dispersion, pooling dispersion information across multiple categories may be an adaptive strategy to increase precision and reduce bias. This adaptation is consistent with using a pooled error term in tests of statistical inference, and more generally with the statistical assumption that the variance from one group contains information about the variance of the other (Keppel and Wickens 2004).

There is also a difference in statistical complexity. Variance is defined as the average squared deviation from the mean. While computing the variance requires knowledge about the mean, computing the mean does not require knowledge about the variance. One might argue that there are shortcuts to estimating the variance that do not require the mean as an input. For instance, because the range of a normal distribution is roughly equal to four times the standard deviation, keeping track of the minimum and maximum values observed would allow people to approximate the standard deviation. However, our studies also demonstrated a spillover effect for the range.

The difference between central tendency and dispersion thus goes beyond statistical complexity. Just like central tendency is the “first statistical moment” and dispersion the “second statistical moment,” central tendency may be the “first intuitive statistic” and dispersion the “second intuitive statistic.”

**Future Research Directions**

We hope our paper may stimulate researchers to further examine the boundary conditions of dispersion spillover. We offer three suggestions. First, our studies examined dispersion spillover for a single attribute learned across multiple categories, but dispersion spillover might occur also across attributes (e.g., prices and battery life) within a category (e.g., phones). Dispersion spillover across attributes, if observed, would suggest that dispersion spillover is a relatively automatic feature of information processing, and that it also happens when the two distributions are ostensibly unrelated.

Second, our studies always included a learning phase and a test phase, but it is not clear if the learning phase is necessary for dispersion spillover to emerge. This depends on whether the spillover happens at the time of encoding, judgment, or both. It is conceivable that people's impressions of dispersion for a given category depend on the amount of dispersion in another category that is activated at the time of judgment, even if they were not encoded at the same time. This could give rise to interesting framing effects. For instance, consumers might ascribe more price dispersion to Gucci handbags when they are evaluated in the context of "luxury goods" (i.e., a more dispersed category) versus "luxury handbags" (a less dispersed category).

Third, our studies have examined participants' ability to form category-specific representations of central tendencies and dispersion, and observed a between-category spillover of dispersion. Future studies could explore the presence of a similar spillover in judgments of covariation (between price and quality for instance). Based on the results of this manuscript, we expect that consumers would also have difficulty forming category-specific impressions of correlation/covariation, and that the strength of association between two variables in one category would be influenced by the correlation of the same two variables in another category.

Finally, we hope future research will explore the role of dispersion spillover in empirical models of consumer learning. For instance, Erdem and Keane's dynamic learning model (1996) assumes that consumers' beliefs about the quality of a brand are not influenced by other brands also present in the learning environment. Our data suggest that the perceived variability of a brand’s quality will be influenced also by the variability in the quality of other brands.

### Distributing Distribution Builders

Distribution builders are powerful tools, and we believe they can be useful to study a wide variety of research questions, for instance in the domains of reference price formation (Mazumdar, Raj, and Sinha 2005) or in service and product quality expectations (e.g., Steenkamp 1990). We hope that the distBuilder library we have created will encourage researchers in academia and business to measure consumer knowledge beyond averages. It is available here: [link removed to keep peer-review blind].

## DATA COLLECTION INFORMATION

Data for all studies were collected by the first author on the following dates:[6]

- Study 1 started on March 13, 2019 at 20:45 PM and ended on March 14, 2019 at 01:01 AM.
- Study 2 started on August 05, 2019 at 22:16 PM and ended on August 06, 2019 at 01:07 AM.
- Study 3 started on February 07, 2020 at 16:59 PM and ended on February 07, 2020 at 18:58 PM.
- Study 4 started on June 27, 2019 at 14:14 PM and ended on June 27, 2019 at 16:01 PM.

[6] All day and times are expressed in UTC.

- Study 5 was preregistered on February 15, 2020 at 16:31 PM, started on February 15, 2020 at 16:52 PM and ended on February 16, 2020 at 02:07 AM.
- Study 6 was preregistered on July 16, 2019 at 19:05 PM, started on July 16, 2019 at 19:47 PM and ended on July 17, 2019 at 01:29 AM.
- Study 7 was preregistered on March 27, 2020 at 13:28 PM. Data collection for the online sample started on March 27, 2020 at 14:37 PM and ended on March 29, 2020 at 07:44 AM. Data collection for the lab samples started on April 08, 2020 at 13:04 PM and ended on April 28, 2020 at 08:06 AM.
- Study 8 started on February 22, 2017 at 17:47 PM and ended on February 22, 2017 at 20:43 PM.
- Study 9 started on August 07, 2017 at 21:55 PM and ended on August 08, 2017 at 01:21 AM.
- Study 10 started on February 06, 2017 at 18:51 PM and ended on February 07, 2017 at 00:34 AM.

The first author is responsible for all the statistical analysis and graphs reported in this manuscript and in the online appendix. The experimental materials, raw data, codebooks, data transformation scripts, and data analysis scripts for all studies have been posted on the Open Science Framework (OSF) repository of the paper.

## REFERENCES

Beach, Lee Roy and Thomas S. Scopp (1968), “Intuitive Statistical Inferences about Variances,” *Organizational Behavior and Human Performance*, 3(2), 109–23.

Bearden, J. Neil and Ryan O. Murphy (2007), “On Generalized Secretary Problems,” in *Uncertainty and Risk*, ed. Mohammed Abdellaoui, R. Duncan Luce, Mark J. Machina, and Bertrand Munier, Berlin, Heidelberg: Springer Berlin Heidelberg, 187–205, http://link.springer.com/10.1007/978-3-540-48935-1_11.

Brehmer, Berndt (1973), “Single-Cue Probability Learning as a Function of the Sign and Magnitude of the Correlation between Cue and Criterion,” *Organizational Behavior and Human Performance*, 9(3), 377–95.

Brehmer, Berndt and Lars-åke Lindberg (1970), “The Relation between Cue Dependency and Cue Validity in Single-Cue Probability Learning with Scaled Cue and Criterion Variables,” *Organizational Behavior and Human Performance*, 5(6), 542–54.

Casasanto, Daniel (2008), “Similarity and Proximity: When Does Close in Space Mean Close in Mind?,” *Memory & Cognition*, 36(6), 1047–56.

Chong, Sang Chul and Anne Treisman (2005), “Statistical Processing: Computing the Average Size in Perceptual Groups,” *Vision Research*, 45(7), 891–900.

Crawford, Vincent P. and Juanjuan Meng (2011), “New York City Cab Drivers’ Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income,” *American Economic Review*, 101(5), 1912–32.

Crosetto, Paolo, Antonio Filippin, Peter Katuščák, and John Smith (2020), “Central Tendency Bias in Belief Elicitation,” *Journal of Economic Psychology*, 78, 102273.

DeLosh, Edward L., Jerome R. Busemeyer, and Mark A. McDaniel (1997), “Extrapolation: The Sine qua Non for Abstraction in Function Learning.,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 968.

Dietvorst, Berkeley and Soaham Bharti (2019), *People Reject Even the Best Possible Algorithm in Uncertain Decision Domains*, SSRN Scholarly Paper ID 3424158, Rochester, NY: Social Science Research Network, https://papers.ssrn.com/abstract=3424158.

Erdem, Tülin and Michael P. Keane (1996), “Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets,” *Marketing science*, 15(1), 1–20.

Flannagan, Michael J., Lisbeth S. Fried, and Keith J. Holyoak (1986), “Distributional Expectations and the Induction of Category Structure,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(2), 241–56.

Fried, Lisbeth S. and Keith J. Holyoak (1984), “Induction of Category Distributions: A Framework for Classification Learning,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(2), 234–57.

Gallagher, Emily, Radhakrishnan Gopalan, Michal Grinstein-Weiss, and Jorge Sabat (2018), “Medicaid and Household Savings Behavior: New Evidence from Tax Refunds.”

Gigerenzer, Gerd and David J. Murray (2015), *Cognition as Intuitive Statistics*, Psychology Press.

Gigerenzer, Gerd and Reinhard Selten (2002), *Bounded Rationality: The Adaptive Toolbox*, MIT press.

Goldstein, Daniel G., Eric J. Johnson, and William F. Sharpe (2008), “Choosing Outcomes versus Choosing Products: Consumer-Focused Retirement Investment Advice,” *Journal of Consumer Research*, 35(3), 440–56.

Goldstein, Daniel G. and David Rothschild (2014), “Lay Understanding of Probability Distributions,” *Judgment and Decision Making*, 9(1), 1.

Goldstein, Daniel G. and Nassim Nicholas Taleb (2007), “We Don’t Quite Know What We Are Talking About,” *The Journal of Portfolio Management*, 33(4), 84–86.

Hahn, Ulrike and Nick Chater (1998), “Similarity and Rules: Distinct? Exhaustive? Empirically Distinguishable?,” *Cognition*, 65(2), 197–230.

Hampton, James A. (1979), “Polymorphous Concepts in Semantic Memory,” *Journal of Verbal Learning and Verbal Behavior*, 18(4), 441–61.

Juslin, Peter, Henrik Olsson, and Anna-Carin Olsson (2003), “Exemplar Effects in Categorization and Multiple-Cue Judgment.,” *Journal of Experimental Psychology: General*, 132(1), 133.

Kalwani, Manohar U., Chi Kin Yim, Heikki J. Rinne, and Yoshi Sugita (1990), “A Price Expectations Model of Customer Brand Choice,” *Journal of Marketing research*, 251–62.

Kareev, Yaakov, Sharon Arnon, and Reut Horwitz-Zeliger (2002), “On the Misperception of Variability.,” *Journal of Experimental Psychology: General*, 131(2), 287.

Keppel, Geoffrey and Thomas D. Wickens (2004), *Design and Analysis: A Researcher’s Handbook*, 4 edition, Upper Saddle River, N.J: Pearson.

Laestadius, John E. Jr. (1970), “Tolerance for Errors in Intuitive Mean Estimations,” *Organizational Behavior and Human Performance*, 5(2), 121–24.

de Langhe, Bart (2016), “The Marketing Manager as an Intuitive Statistician,” *Journal of Marketing Behavior*, 2(2–3), 101–27.

Levin, Irwin P. (1974), "Averaging Processes in Ratings and Choices Based on Numerical Information," *Memory & Cognition*, 2(4), 786–90.

——— (1975), "Information Integration in Numerical Judgments and Decision Processes.," *Journal of Experimental Psychology: General*, 104(1), 39.

Lindskog, Marcus, Anders Winman, and Peter Juslin (2013), "Calculate or Wait: Is Man an Eager or a Lazy Intuitive Statistician?," *Journal of Cognitive Psychology*, 25(8), 994–1014.

Long, Andrew R., Philip M. Fernbach, and Bart De Langhe (2018), "Circle of Incompetence: Sense of Understanding as an Improper Guide to Investment Risk," *Journal of Marketing Research*, 55(4), 474–88.

Lovie, Patricia (1978), "Teaching Intuitive Statistics II. Aiding the Estimation of Standard Deviations," *International Journal of Mathematical Education in Science and Technology*, 9(2), 213–19.

Lovie, Patricia and A. D. Lovie (1976), "Teaching Intuitive Statistics I: Estimating Means and Variances," *International Journal of Mathematical Educational in Science and Technology*, 7(1), 29–39.

Malmi, Robert A. and David J. Samson (1983), "Intuitive Averaging of Categorized Numerical Stimuli," *Journal of Memory and Language*, 22(5), 547.

March, James G. (1996), "Learning to Be Risk Averse.," *Psychological Review*, 103(2), 309.

Markowitz, Harry (1952), "The Utility of Wealth," *Journal of political Economy*, 60(2), 151–58.

Mazumdar, Tridib, Sevilimedu P. Raj, and Indrajit Sinha (2005), "Reference Price Research: Review and Propositions," *Journal of marketing*, 69(4), 84–102.

McCall, John J. (1972), "The Simple Mathematics of Information, Job Search, and Prejudice," *Racial discrimination in Economic Life, Lexington Books*, 205–24.

Medin, Douglas L. and Marguerite M. Schaffer (1978), "Context Theory of Classification Learning," *Psychological Review*, 85(3), 207–38.

Mellers, Barbara (1986), "Test of a Distributional Theory of Intuitive Numerical Prediction," *Organizational Behavior and Human Decision Processes*, 38(3), 279–94.

Mellers, Barbara A., Virginia Richards, and Michael H. Birnbaum (1992), "Distributional Theories of Impression Formation," *Organizational Behavior and Human Decision Processes*, 51(3), 313–43.

Page, Lionel and Daniel G. Goldstein (2016), "Subjective Beliefs about the Income Distribution and Preferences for Redistribution," *Social Choice and Welfare*, 47(1), 25–61.

Payne, John W., James R. Bettman, and Eric J. Johnson (1997), *The Adaptive Decision Maker: Effort and Accuracy in Choice*, Cambridge University Press: Cambridge, UK.

Peterson, Cameron and Alan Miller (1964), "Mode, Median, and Mean as Optimal Strategies," *Journal of Experimental Psychology*, 68(4), 363–67.

Peterson, Cameron R. and Lee R. Beach (1967), "Man as an Intuitive Statistician.," *Psychological Bulletin*, 68(1), 29.

Reber, Arthur S. (1989), "Implicit Learning and Tacit Knowledge.," *Journal of experimental psychology: General*, 118(3), 219.

Reinholtz, Nicholas (2015), *Persistence in Consumer Search*, Columbia University.

Reinholtz, Nicholas, Philip Fernbach, and Bart De Langhe (2016), "Do People Understand the Benefit of Diversification?"

Rips, Lance J. (1989), "Similarity, Typicality, and Categorization," *Similarity and Analogical Reasoning*, 2159.

Rosch, Eleanor and Carolyn B Mervis (1975), "Family Resemblances: Studies in the Internal Structure of Categories," *Cognitive Psychology*, 7(4), 573–605.

Rothschild, Michael (1974), "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9(2), 185–202.

Sakamoto, Yasuaki, Matt Jones, and Bradley C. Love (2008), "Putting the Psychology Back into Psychological Models: Mechanistic versus Rational Approaches," *Memory & Cognition*, 36(6), 1057–65.

Shanks, David R. and Mark F. St John (1994), "Characteristics of Dissociable Human Learning Systems," *Behavioral and brain sciences*, 17(3), 367–95.

Simonsohn, Uri (2014), "[17] No-Way Interactions," *Data Colada*, http://datacolada.org/17.

Simonsohn, Uri, Joe Simmons, and Leif Nelson (2015), "AsPredicted," *AsPredicted*, https://aspredicted.org/.

Skinner, Jonathan (1988), "Risky Income, Life Cycle Consumption, and Precautionary Savings," *Journal of Monetary Economics*, 22(2), 237–55.

Smith, Andrew R. and Paul C. Price (2010), "Sample Size Bias in the Estimation of Means," *Psychonomic Bulletin & Review*, 17(4), 499–503.

Smith, Edward E. and Douglas L. Medin (1981), *Categories and Concepts*, 9, Harvard University Press Cambridge, MA.

Spencer, J. (1963), "A Further Study of Estimating Averages," *Ergonomics*, 6(3), 255–65.

Steenkamp, Jan-Benedict EM (1990), “Conceptual Model of the Quality Perception Process,” *Journal of Business research*, 21(4), 309–33.

Stewart, Neil and Nick Chater (2002), “The Effect of Category Variability in Perceptual Categorization.,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(5), 893.

Stigler, George J. (1962), “Information in the Labor Market,” *Journal of political economy*, 70(5, Part 2), 94–105.

——— (1961), “The Economics of Information,” *Journal of Political Economy*, 69(3), 213–25.

Verghese, Preeti (2001), “Visual Search and Attention: A Signal Detection Theory Approach,” *Neuron*, 31(4), 523–35.

Weber, Elke U., Sharoni Shafir, and Ann-Renée Blais (2004), “Predicting Risk Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation.,” *Psychological review*, 111(2), 430.

Wilken, Patrick and Wei Ji Ma (2004), “A Detection Theory Account of Change Detection,” *Journal of Vision*, 4(12), 11–11.

Winkler, Robert L. (1970), “Intuitive Bayesian Point Estimation,” *Organizational Behavior and Human Performance*, 5(5), 417–29.

Wolfe, Mary L. (1975), “Distribution Characteristics as Predictors of Error in Intuitive Estimation of Means,” *Psychological Reports*, 36(2), 367–70.