Online Appendix: Digitization and Pre-Purchase Information: The Causal and Welfare Impacts of Reviews and Crowd Ratings

Imke Reimers
Northeastern University

Joel Waldfogel
University of Minnesota, NBER, ZEW

January 24, 2021

A Estimating the Nested Logit Parameter

We infer the degree of substitutability σ using the Nielsen weekly sales data for the top 100, which we have for 2015-2018. We also need a few additional pieces of information, along with an instrumental variables strategy. We describe these in turn.

First, we obtain weekly data on total physical book sales from Publishers Weekly, which reports this in most but not all weeks.[1] We refer to this as Q_t. We have these data for 124 weeks during 2015-2018. First, as above, we assume that each member of the US population is making a monthly decision of whether to purchase a book, so with weekly observations, the market size is $M = 0.25 \times 327$ million.

We then define the following variables:

$$s_{jt} = q_{jt}/M, \qquad s_{jt|g} = q_{jt}/Q_t, \qquad s_{0t} = 1 - Q_t/M.$$

As in Berry (1994), we seek to obtain σ from a regression of $\ln(s_j) - \ln(s_0)$ on $\ln(s_{jt|g})$. Intuitively, identification comes from the relationship between the number of products available and whether the share of the population buying books increases.

The estimation requires some additional thought. First, there is seasonality in the book market, with a substantial increase in sales around Christmas. Publishers know this and may release more books around Christmas, raising a concern that the number of books coming out as well as demand might rise around Christmas. This would look like an effect of product entry on market expansion, even if it were not. To address this, we include week-of-the-year dummies.

Second, we need an instrument for the books' inside shares $s_{jt|g}$. One natural idea would be the number of products available in each week. In our data it is by construction 100. More to the point, however, not all products are of equivalent importance. We appeal to the logic of BLP instruments, which are terms involving the other products in the choice set. Here, for example, we measure the number of products in the top 100 that were originally released in the past week, in the past two weeks, and so on, up to ten weeks. Further, because we have the Nielsen weekly top 100 going back to 2015, we construct measures of authors' past sales. We then use measures of the past sales of authors whose new books are in the top 100 this week. We implement this with the number of authors in the current top 100 whose previous sales are in one of seven intervals.

This gives us 17 possible instruments – ten representing the books' "vintages," and seven representing the authors' histories. To avoid choosing among them arbitrarily, we use the variable selection approach of Belloni, Chernozhukov and Hansen (2014). We estimate IV regressions in which we use LASSO techniques for the choices of a) which week dummies to include in the main equation, and b) which instruments to include in the first stage. The procedure selects 4 of the 17 possible instruments and 16 of the possible week dummies. Not surprisingly, the weeks before Christmas are selected. The resulting estimate of σ is 0.373 (se=0.0557).

[1] See Publishers Weekly Editors (2017–2019).

B Additional Robustness of Welfare Results

B.1 Price and B Parameters

The price parameter α determines the absolute size of a welfare effect. It does not, however, affect the relative sizes of the respective effects of professional reviews and crowd ratings on consumer surplus, so our conclusions on the relative impacts of professional reviews and star ratings are unaffected by α.

The elasticity of the quantity sold with respect to the sales rank can affect our results. The value we use for this elasticity is drawn from our B estimate (0.47), which is similar to both estimates we obtain from weekly data as well as other research. Using annual data on book sales quantities and ranks, Liebowitz and Zentner (2020) find roughly 0.5 for the top 2,000 titles and 0.6 for the next 8,000, and larger absolute-value elasticities farther down the distribution. Using a higher estimate of the sales quantity-rank elasticity increases the effect of stars on CS relative to the effect of professional reviews because professionally reviewed books have, on average, better sales ranks. For example, when we use twice our estimate of B for the elasticities in equations (**??**) and (**??**), we obtain an estimate of the ΔCS from stars of 80.51 and an analogous estimate for reviews of 1.23. The resulting ratio is 65.45. Hence, our result – that stars have a larger effect than reviews – is not driven by our chosen estimate of the elasticity of sales quantities with respect to ranks.

B.2 Richer Nesting Structure

Our baseline (one-level) nested logit model relies on the logit error as the only reason that different consumers choose different books. We can calculate ΔCS using a more flexible two-level nested logit model in which consumers have different preferences for genres: First, consumers choose whether to purchase books or the outside good. They then choose a genre, and then a book within a genre. This approach allows for horizontal differentiation, in that books can be closer substitutes within genres than across. We calculate the changes in consumer surplus with a range of values of two substitution parameters, σ_1, which reflects the degree of substitutability across titles within a genre, and σ_2, which reflects the substitutability of genres for one another, maintaining that books within a genre are closer substitutes than different genres ($\sigma_1 > \sigma_2$). Using σ_2 from 0 to .95 and σ_1 from σ_2 to .95, we find that ΔCS from stars varies between \$34.67 and \$36.58 million, while ΔCS from reviews varies between \$2.54 and \$3.27 million. The ratio of the effects varies from 11.18 to 13.99. As in the simpler nested logit model, the results remain quite similar to the baseline results across a range of substitution patterns. We conclude that our results are not driven by the particular substitution patterns embodied in the baseline model.

B.3 Sample Representativeness

Our sample includes a comprehensive list of weekly USA Today top 150 bestsellers but not all books in the market, so our sample over-represents titles at the head of the sales distribution. When we scale the sample estimate of the change in CS from stars up to the population, we implicitly assume that the change in CS per title would be similar for average books in the population. We can instead take a more conservative approach that only scales up the ΔCS from books outside the bestseller lists.

Using our translation of ranks to quantities, total sample sales are 331 million; and USA Today bestsellers account for 219 million of these. If we decompose the change in CS between bestsellers and other books, we find that the 219 million bestsellers sold deliver \$27.7 million in increased CS from stars (or \$0.127 per unit), while the remaining 112 million in sample sales account for \$8.11 million in increased CS (\$0.072 per unit). This leads almost exactly – with the approximation arising only from the nonlinearity of the logit – to the baseline ΔCS estimate of \$35.83 from stars. We can re-weight the per-book ΔCS estimates as if bestsellers made up 31.5 percent (=219/695) of total sales, while non-bestsellers account for the remaining share. This delivers a ΔCS of \$27.63 million rather than our baseline of \$35.83 million, and the ΔCS ratio becomes 8.69, compared to the baseline 11.27. We conclude that our main result – that ΔCS from stars substantially exceeds ΔCS from professional reviews – remains even when accounting for possible representativeness issues in the sample.

References

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. “High-dimensional methods and inference on structural and treatment effects.” *Journal of Economic Perspectives*, 28(2): 29–50.

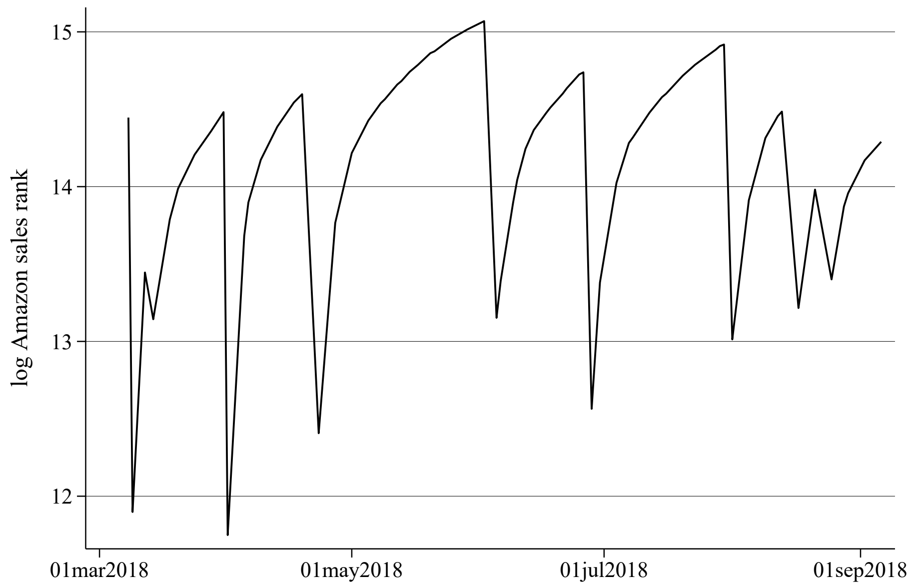
Berry, Steven T. 1994. “Estimating discrete-choice models of product differentiation.” *The RAND Journal of Economics*, 242–262.

Liebowitz, Stan J, and Alejandro Zentner. 2020. “The Challenges of Using Ranks to Estimate Sales.” *Available at SSRN 3543827.*

Publishers Weekly Editors. 2017–2019. “Weekly Scorecard, various issues.” *Publishers Weekly.*

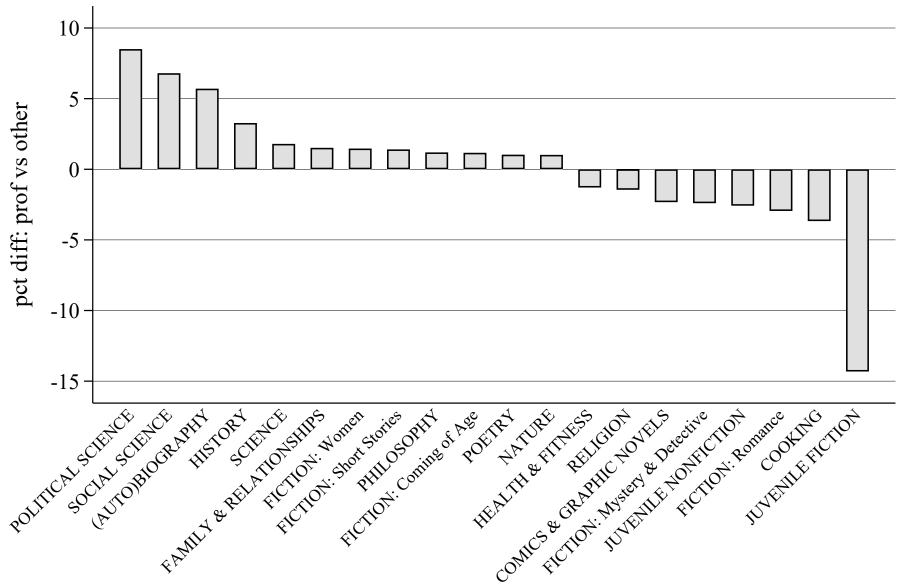
C Additional Figures

Figure 1: Amazon sales rank evolution of a sample book



Notes: Daily Amazon US log sales rank data for Amazon ASIN 198210029X. Large improvements show a day with a sale, followed by days of upward drift until the next sale.

Figure 2: Composition of genres – reviewed vs. not



Notes: This figure reports the difference in the genre distributions between the professional outlets and others: the share of each genre in the professionally reviewed sample minus the share of the same genre in the remaining books. We include only the genres that differ by at least one percentage point.