Having an Ambiguous Gender

Dana Beth Weinberg*[1] and Adam Kapelner†[2]

[1]*Dept. of Sociology, Queens College, The City University of New York*
[2]*Dept. of Mathematics, Queens College, The City University of New York*

April 13, 2016

Abstract

When analyzing survey and experimental data, social scientists generally consider only gender being binary: male and female. An ambiguous gender is generally treated as a nuisance which needs to be imputed to either male or female or those rows omitted from the analysis. We demonstrate here that ambiguous or "androgynous" gender is an important third category that should be considered for future research. We first observe in a dataset of published book prices that authors whose names are of androgynous gender (first name being initials) command higher book prices. We could not rule out omitted variable bias when assessing if this effect was causal. So we embarked on a large $n \approx 2500$ randomized natural field experiment using Amazon's Mechanical Turk (we discuss this methodology herein in detail). We provided cover images, titles and description blurbs from published e-books in two genres while manipulating the author's gender: male, female and androgynous. We asked participants to assess the book's value and found that people would indeed pay more ($0.25) for an androgynously-authored book. We connect this result to the psychology of suprise literature and discuss implications.

*Electronic address: `Dana.Weinberg@qc.cuny.edu`; Prinicipal Corresponding author
†Electronic address: `Kapelner@qc.cuny.edu`; Corresponding author

1 Introduction and Motivation

Consider the following scenario facing us: we were analyzing book publishing data for gender bias from xx containing every book published in the past yy years with information about price, author(s), year, genre, etc. In order to assess gender bias, we first need the gender of the author. So we use a lookup table of common male and female names and match based on a threshold score. We code multiple authors as "ineligible" and we code only initials as "initials" and everything else as "A".

Our gender variable looks like this:

```
M   F   A       ineligible    initials
833 848 863     1111          333
```

To investigate bias, we would regress a dummy for male (or female) on price and attempt to control for all omitted variable bias using other covariates in the dataset (genre, year, publisher information, etc). In this case, we would drop all non-male or non-female

where

The rest of this paper is organized as follows. We examine the background literature in Section 2. We describe experimentation using MTurk in Section 3. Our experimental design is found in Section 4. We discuss our results in Section 5 and conclude and offer future directions in Section 6.

2 Background Research

Foschi (1996) Dixon et al. (2015) Cheng et al. (2011) Bortolussi et al. (2010) Borsuk et al. (2009) Bobbitt-Zeher (2011) Barnes (2015) Alexander and Bowler (2014) Alexander and Andersen (1993) Ridgeway (1997) Yampbell (2005) van Dijk (2014) van den Brink and Benschop (2012) Uscinski and Goren (2011) Tregenza (2002) Top (1991) Paludi and Bauer (1983) Lloyd (1990) Lena and Lindemann (2014) Leemans and Stokmans (1992) Johnston et al. (2014) Haswell and Haswell (1996) Gorman (2005) Fulton (2012) Foschi (2000)

3 Experimentation on MTurk

Internal Validity

Yes

External Validity

Yes... external validity is only a problem

MTurk's potential as a platform for field experimentation using the framework proposed in Levitt and List (2007, 2009) is explored in Chandler and Kapelner

Amazon's Mechanical Turk (**?**) is the largest online, task-based labor market and is used by hundreds of thousands (**?**). of people worldwide. Individuals and companies can post tasks (known as Human Intelligence Tasks, or "HITs") and have them completed by an on-demand labor force. Typical tasks include image labeling, audio transcription, and basic Internet

research. Academics also use MTurk to outsource low-skilled resource tasks such as identifying linguistic patterns in text (Sprouse, 2011) and labeling medical images (Holmes and Kapelner, 2010). The image labeling system from the latter study, known as "DistributeEyes," was originally used by breast cancer researchers and was modified for our experiment. Beyond simply using MTurk as a source of labor, academics have also began using MTurk as a way to conduct online experiments. The remainder of the section highlights some of the ways this subject pool is used and places special emphasis on the suitability of the environment for natural field experiments in economics.

Berinsky et al. (2012) Buhrmester et al. (2011) Henrich et al. (2010) Paolacci et al. (2010) Gneezy and List (2006) Levitt and List (2007) Levitt and List (2009) Kapelner and Chandler (2010) Chandler and Kapelner (2013) Horton et al. (2011) Harrison and List (2004)

4 Experimental Design

5 Results

6 Discussion

Replication

The experiment performed herein can be duplicated with the Ruby-on-Rails code found in the github repository at `https://github.com/kapelner/neurotika_experiment`. The figures and tables in this manuscript can be replicated by running `paper_duplication.R` found in the root of the repository. All code and scripts are open source under the MIT license.

Acknowledgements

We thank Rikki Katz for help with software engineering for the experiment and Marie Le Pichon for the artwork used in the experiment.

References

Alexander, D. and Andersen, K. (1993). Gender as a Factor in the Attribution of Leadership Traits. *Political Research Quarterly*, 46(3):527–545.

Alexander, V. D. and Bowler, A. E. (2014). Art at the crossroads: The arts in society and the sociology of art. *Poetics*, 43:1–19.

Barnes, J. L. (2015). Fanfiction as imaginary play: What fan-written stories can tell us about the cognitive science of fiction. *Poetics*, 48:69–82.

Berinsky, A. J., Huber, G. A., and Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, 20:351–368.

Bobbitt-Zeher, D. (2011). Gender Discrimination at Work: Connecting Gender Stereotypes, Institutional Policies, and Gender Composition of Workplace.

Borsuk, R. M., Aarssen, L. W., Budden, A. E., Koricheva, J., Leimu, R., Tregenza, T., and Lortie, C. J. (2009). To Name or Not to Name: The Effect of Changing Author Gender on Peer Review. *BioScience*, 59(11):985–989.

Bortolussi, M., Dixon, P., and Sopcák, P. (2010). Gender and reading. *Poetics*, 38(3):299–318.

Buhrmester, M., Kwang, T., and Gosling, S. (2011). Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1):3–5.

Chandler, D. and Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization*, 90:123–133.

Cheng, N., Chandramouli, R., and Subbalakshmi, K. P. (2011). Author gender identification from text. *Digital Investigation*, 8(1):78–88.

Dixon, P., Bortolussi, M., and Sopcák, P. (2015). Extratextual effects on the evaluation of narrative texts. *Poetics*, 48:42–54.

Foschi, M. (1996). Double Standards in the Evaluation of Men and Women. *Social Psychology Quarterly*, 59(3):237–254.

Foschi, M. (2000). Double Standards for Competence: Theory and Research. *Annual Review of Sociology*, 26:21–42.

Fulton, S. A. (2012). Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success. *Political Research Quarterly*, 65(2):303–314.

Gneezy, U. and List, J. A. (2006). Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments. *Econometrica*, 74(5):1365–1384.

Gorman, E. H. (2005). Gender Stereotypes, Same-Gender Preferences, and Organizational Variation in the Hiring of Women: Evidence from Law Firms. *American Sociological Review*, 70(4):702–728.

Harrison, G. W. and List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, 42(4):1009 – 1055.

Haswell, R. H. and Haswell, J. T. (1996). Gender Bias and Critique of Student Writing. *Assessing Writing*, 3(1):31–83.

Henrich, J., Heine, S., and Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3):61–83.

Horton, J. J., Rand, D. G., and Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, 14:399–425.

Johnston, J., Rodney, A., and Chong, P. (2014). Making change in the kitchen? A study of celebrity cookbooks, culinary personas, and inequality. *Poetics*, 47:1–22.

Kapelner, A. and Chandler, D. (2010). Preventing Satisficing in Online Surveys: A "Kapcha" to Ensure Higher Quality Data. In *CrowdConf 2010*.

Leemans, H. and Stokmans, M. (1992). A descriptive model of the decision making process of buyers of books. *Journal of Cultural Economics*, 16(2):25–50.

Lena, J. C. and Lindemann, D. J. (2014). Who is an artist? New data for an old question. *Poetics*, 43:70–85.

Levitt, S. and List, J. (2009). Field experiments in economics: the past, the present, and the future. *European Economic Review*, 53(1):1–18.

Levitt, S. D. and List, J. A. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives*, 21(2):153–174.

Lloyd, M. E. (1990). Gender factors in reviewer recommendations for manuscript publication. *Journal of applied behavior analysis*, 23(4):539–543.

Paludi, M. a. and Bauer, W. D. (1983). Goldberg Revisited : What's in an Author's Name. *Sex Roles*, 9(3):387–390.

Paolacci, G., Chandler, J., and Ipeirotis, P. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5):411–419.

Ridgeway, C. L. (1997). Interaction and the Conservation of Gender Inequality: Considering Employment. *American Sociological Review*, 62(2):218–235.

Top, T. J. (1991). Sex bias in the evaluation of performance in the scientific, artistic, and literary professions: A review. *Sex Roles*, 24(1/2):73–106.

Tregenza, T. (2002). Gender bias in the refereeing process? *Trends in Ecology and Evolution*, 17(8):349–350.

Uscinski, J. E. and Goren, L. J. (2011). What's in a Name? Coverage of Senator Hillary Clinton during the 2008 Democratic Primary. *Political Research Quarterly*, 64(4):884–896.

van den Brink, M. and Benschop, Y. (2012). Gender practices in the construction of academic excellence: Sheep with five legs. *Organization*, 19(4):507–524.

van Dijk, Y. (2014). Amateurs online: Creativity in a community. *Poetics*, 43:86–101.

Yampbell, C. (2005). Judging a Book by Its Cover: Publishing Trends in Young Adult Literature. *The Lion and the Unicorn*, 29(3):348–372.

A Demographic Questions

B Response Questions