

Does Diversity Pay? A Replication of Herring (2009)

American Sociological Review
1–11
© American Sociological
Association 2017
DOI: 10.1177/0003122417714422
journals.sagepub.com/home/asr



Dragana Stojmenovska,^a Thijs Bol,^a
and Thomas Leopold^a

Abstract

In an influential article published in the *American Sociological Review* in 2009, Herring finds that diverse workforces are beneficial for business. His analysis supports seven out of eight hypotheses on the positive effects of gender and racial diversity on sales revenue, number of customers, perceived relative market share, and perceived relative profitability. This comment points out that Herring's analysis contains two errors. First, missing codes on the outcome variables are treated as substantive codes. Second, two control variables—company size and establishment size—are highly skewed, and this skew obscures their positive associations with the predictor and outcome variables. We replicate Herring's analysis correcting for both errors. The findings support only one of the original eight hypotheses, suggesting that diversity is nonconsequential, rather than beneficial, to business success.

Keywords

replication, diversity, gender, race, organizations

In an *American Sociological Review* article, Herring (2009:208) makes a “business case for diversity.” Using a sample from the National Organizations Survey (NOS) of 1996 and 1997 (Kalleberg, Knoke, and Marsden 2001), Herring finds that organizations with higher levels of gender and racial diversity are more successful in terms of (1) sales revenue, (2) number of customers, (3) perceived relative market share, and (4) perceived relative profitability. The study supports seven out of eight hypotheses on the positive associations of gender and racial diversity with each of these outcomes. Based on his findings, Herring concludes that “the multivariate analyses strongly support the business case for diversity perspective” (p. 219). Herring's article has become a standard reference in research on the effects of organizational diversity, as evidenced by 522 citations in Google Scholar and 142 citations in the Web of Science (as of May 18, 2017).

In this comment, we show that Herring's analyses contain two errors.¹ First, he treated missing codes on the outcome variables as substantive codes. Depending on the outcome variable, a correct specification of missing codes leads to analytic samples that are between 53 and 25 percent smaller than Herring's. Second, Herring did not adequately control for two important confounders: company size and establishment size. In Herring's analysis, both variables are highly skewed, and this skew obscures their positive associations with the predictor variables and the outcome variables. The skewed control variables used

^aUniversity of Amsterdam

Corresponding Author:

Thomas Leopold, University of Amsterdam,
Nieuwe Achtergracht 166, 1018WV Amsterdam,
The Netherlands
E-mail: t.leopold@uva.nl

in Herring's analyses violate the linearity assumption of OLS regression, resulting in underestimation of the true relationship between the independent variables and the dependent variable.²

In the following sections, we first describe both errors and then replicate Herring's analysis correcting for both errors. The results from our replication support only one of the original eight hypotheses.

Apart from correcting Herring's analysis, this comment adds to the literature in two more general ways. First, we contribute to the current replication debate, which has largely centered on instances of scientific misconduct (e.g., Broockman, Kalla, and Aronow 2015). Our replication directs attention to a less prominent but more prevalent issue—ordinary errors—and the value of independent verification of research results (Firebaugh 2007; Freese 2007). We highlight the need for shared professional standards for documenting quantitative empirical work (e.g., in the form of replication packages) to prevent and discover errors that occur in the research process.

Second, we demonstrate the risk of bias associated with skewed variables that violate the assumptions of OLS regression—a risk that is commonly discussed in the methodological literature and textbooks (e.g., Allison 1999; Berry 1993; Fox 1991) but often ignored in research (Osborne, Christianson, and Gunter 2001). In replicating Herring's analysis, we document a case in which conclusions change after skewed variables are transformed.

MISSING VALUES AND SAMPLE SIZES

Herring's analytic sample consists of "506 for-profit business organizations that provided information about the racial composition of their full-time workforces, their sales revenue, their number of customers, their market share, and their profitability" (p. 213). In our replication we were unable to recover this sample. Table 1 illustrates the problem.

The right-hand column shows Herring's sample sizes for each outcome, taken from the tables in which he regresses each outcome

first only on the two diversity indicators³ and then on the diversity indicators and a vector of control variables (Tables 2 and 3 in Herring [2009]). The middle column shows the number of cases left after excluding business organizations with missing values on the outcome and predictor variables. A comparison between the middle and right-hand columns shows that Herring's sample sizes considerably exceed the number of valid cases available on for-profit business organizations in the NOS. The discrepancies are particularly large for sales revenues and the number of customers, and smaller for perceived relative market share and perceived relative profitability.

In our correspondence with Herring, he did not offer a definitive explanation for these discrepancies, but indicated that he may have treated all codes other than "not applicable" (−999) as substantive codes. Given (1) the large difference between his sample size and the number of valid observations in the NOS, and (2) the large number of missing values due to reasons other than "not applicable"—in particular for sales revenue and number of customers—this coding error appears likely to account for much of the discrepancies. This means, for example, that 206 business organizations in which the sales revenue was unknown were treated as if they had sales of 88,888,888,888 US Dollars. Yet, even when we replicated this error (i.e., keeping all organizations with missing values other than −999 in our sample), we were unable to recover Herring's sample sizes, although the differences were smaller.

Alternative explanations such as (1) increasing the sample size by pooling the 2002 round of the NOS, (2) use of imputation techniques, and (3) use of flag variables for missing values on the predictors are also inconsistent with the sample sizes reported by Herring. Moreover, the original article does not indicate that any such techniques were used. Another possibility is that Herring did not correctly report his sample sizes in this article. The increase in the sample size after adding covariates in the models on perceived relative profitability (bottom lines of Table 1) points to errors in reporting.

Table 1. Missing Data in the NOS 1996/97 for the Four Dependent Variables

	Cases Lost	Cases Left	Herring's Count
Sales ^a		710	
Missing data on outcome			
Not applicable (–999)	58	652	
Don't know (88,888,888,888)	206	446	
Refused to answer (99,999,999,999)	64	382	
+ Missing data on diversity indicators	55	327	506
+ Missing data on control variables	88	239	506
Customers ^b		710	
Missing data on outcome			
Not applicable (–999)	60	650	
Don't know (88,888,888)	224	426	
Refused to answer (99,999,999)	7	419	
+ Missing data on diversity indicators	55	364	506
+ Missing data on control variables	94	270	506
+ Dropping organizations with 0 customers ^e	4	266	
Market Share ^c		710	
Missing data on outcome			
Not applicable (–999)	42	668	
Don't know (6)	66	602	
Refused to answer (7)	9	593	
+ Missing data on diversity indicators	100	493	506
+ Missing data on control variables	145	348	469
Profitability ^d		710	
Missing data on outcome			
Not applicable (–999)	41	669	
Don't know (6)	49	620	
Refused to answer (7)	8	612	
+ Missing data on diversity indicators	102	510	470
+ Missing data on control variables	148	362	484

Note: National Organizations Survey 1996/97, own calculations. The counts from the original article can be found in Herring's Table 2 (Herring 2009:217) and Table 3 (Herring 2009:218).

^aAnnual sales revenue over the past two years.

^bNumber of customers.

^cPerceived market share relative to other businesses.

^dPerceived profitability relative to other businesses.

^eOrganizations with 0 customers are dropped in the final analyses (Table 3, columns SBL 2) due to the log transformation.

In our correspondence, Herring did not offer further information about his sample selection and treatment of missing values. Hence we were unable to determine the exact causes of the discrepancies between the sample sizes reported by Herring and those found in our replication. Given that Herring's article does not indicate that he deviated from the default technique—listwise deletion—we proceeded in our replication by dropping

from the models all business organizations that had missing values on the outcome variables or predictor variables. As Table 1 shows, listwise deletion yields sample sizes that vary between 362 (perceived relative profitability) and 239 (sales revenue).

Listwise deletion drops all observations with a missing value on at least one of the covariates included in the model. In additional analyses, we replicated the multivariable analyses using

multiple imputation instead of listwise deletion. The code for these analyses is included in the replication package. Multiple imputation yielded results similar to those presented here.

SKEWED VARIABLES

The size of business organizations constitutes a potential confounder in Herring's analyses. If larger organizations are more diverse in terms of their gender composition and racial composition, positive effects of diversity on business outcomes may be spurious, given that the size of an organization is positively related to all four business outcomes. To control for this, Herring includes untransformed count variables for company size and establishment size⁴ in his fully specified models that are used to evaluate the hypotheses (Table 3 in Herring [2009]).

The left-hand panels in Figure 1 show that the distributions of these count variables are highly skewed. The same is true for two of the outcome measures, sales revenue and the number of customers. A common technique to address right-skew is to take the natural logarithm of these variables and then use the transformed variables in the regression. Herring uses the log transformation for only one of the skewed variables—sales revenue—and leaves the remaining three variables untransformed. As illustrated in Figure 1, the log transformation of company size, establishment size, and number of customers eliminates most of the skew.

Transforming skewed variables is not required per se, and the inclusion of skewed variables does not always violate the assumptions of OLS regression or change the results in a meaningful way. Yet in Herring's analysis it does. To illustrate why, Table 2 shows how the untransformed and log-transformed control variables for establishment size and company size are correlated with the predictors and outcomes of interest. The key predictors are gender diversity (AID-G) and racial diversity (AID-R);⁵ the outcomes are sales revenue, number of customers, perceived relative market share, and perceived relative profitability.

The predictors as well as the outcomes are positively associated with the size of a

business organization, but the untransformed measures underestimate these correlations. After logging establishment size and company size, their correlations with racial diversity increase from .09 to .42 and from .14 to .41, respectively; correlations with gender diversity increase from .08 to .15 and from .04 to .17, respectively. Taking the natural logarithm of the size of a business organization also yields higher correlations with the four outcome variables. For example, the correlations of establishment size and company size with sales revenue increase from .34 to .75 and from .31 to .73, respectively; correlations with number of customers increase from .17 to .21 and from .08 to .23, respectively.

These results show that the untransformed controls that Herring used in his analysis fail to pick up much of their correlation with the diversity and outcome measures. As a result, Herring's regression models did not adequately control for the confounding influence of the size of a business organization. The component-plus-residual plots in Figure 2 show how the untransformed measures of establishment size violate the linearity assumption of OLS (Allison 1999).⁶ Log-transforming the variables ensures a better fit with the assumptions of OLS regression.

REPLICATION RESULTS

Table 3 shows the results of our replication of Herring's analysis. We replicate the fully specified models that include all control variables (Table 3 in Herring [2009]). For each outcome variable, we present three columns: the first column shows the results as reported by Herring; the second column shows the results when correcting only for the erroneous specification of missing values; the third column shows the results when additionally correcting for skew by log-transforming establishment size, company size, and number of customers. All remaining controls that we include are also included in Herring's original analyses.⁷

The left-hand column of each outcome variable shows Herring's estimates. Seven out of eight parameter estimates of the diversity indicators are positive and statistically significant.

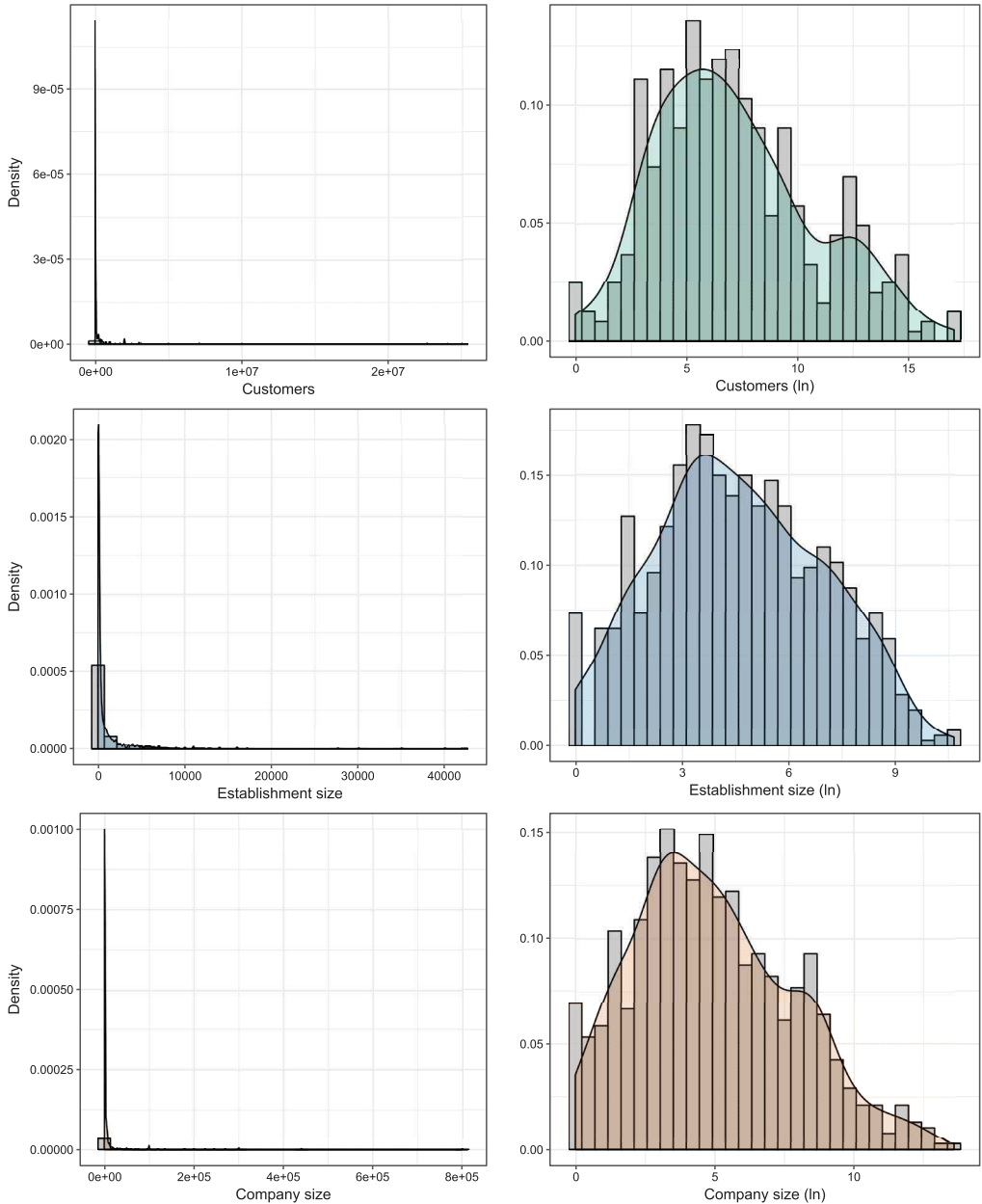


Figure 1. Distributions of the Skewed Count Variables Before and After Logging
Note: NOS 1996/97, own calculations. The left-hand panels display the distributions of the untransformed variables; the right-hand panels display the distributions of the variables that are transformed by taking the natural logarithm.

After correcting for the erroneous missing codes (middle column), all point estimates remain positive, but only three are different from zero at conventional levels of statistical significance ($p < .05$, two-tailed). In the

right-hand column, we additionally correct for skew in three variables (number of customers, establishment size, and company size). In these models, all remaining effects disappear, except for the positive association between gender

Table 2. Pairwise Correlations with Untransformed and Transformed Measures of Establishment Size and Company Size

	Establishment Size		Company Size	
	Linear	Logged	Linear	Logged
Racial Diversity	.094	.416	.138	.411
Gender Diversity	.076	.151	.044	.166
Sales Revenue	.338	.746	.305	.729
Number of Customers	.165	.209	.083	.228
Market Share	.154	.238	.089	.230
Profitability	.117	.116	.002	.121

Note: National Organizations Survey 1996/97, own calculations. All correlations are Pearson's R and are calculated over the 217 businesses for which information on the diversity indicators, the four business outcomes, and establishment and company size is non-missing. Pairwise correlations calculated over the maximum number of observations for that pair yields similar results.

diversity and the number of customers. When evaluating Herring's hypotheses on the basis of statistical significance (as Herring did), this means that only one out of eight hypotheses is supported. In the original article, seven out of eight hypotheses were supported.

Figure 3 illustrates these findings in the form of standardized effects plots. After both errors are corrected, the only statistically significant effect indicates that a one standard deviation increase in gender diversity is associated with a .15 standard deviation increase in the logged number of customers. All remaining parameter estimates for diversity effects are not only insignificant, but also small in size, indicating that the NOS data of 1996/97 yield little evidence to support a business case for diversity.

CONCLUSIONS

In his 2009 *American Sociological Review* article, Herring argued that diversity pays. Based on aggregate organizational data from the NOS, he concluded "that a diverse workforce is good for business, offering a direct return on investment and promising greater corporate profits and earnings" (Herring 2009:220). In this comment, we have shown that Herring's analysis contains two errors. When these errors are corrected, the results no longer support Herring's conclusions.

In the years following its publication, Herring's article has become an influential source

of empirical support for the value-in-diversity perspective. Our replication shows that the data Herring analyzed are not consistent with this perspective. The overall pattern of findings suggests that diversity is nonconsequential, rather than beneficial, to business success.

This comment emerged from a student assignment⁸ in a graduate course on replication offered in the *Research Master Social Sciences* program at the University of Amsterdam. The aim of this course was not to hunt for errors in the work of other scholars, but to give students an opportunity to "learn from the best" (King 2006) by replicating studies published in a discipline's premier journals. As we have seen, even work that is reviewed as intensively as articles submitted to *ASR* can contain serious technical errors that lead to statistical artifacts rather than robust findings.

Much of the current interest in replication focuses on (1) issues of research design (Open Science Collaboration 2015) and (2) cases of scientific misconduct, ranging from the manipulative use of methods to obtain specific (statistically significant) results to outright fraud (e.g., Broockman et al. 2015). This comment directs attention to the importance of replication for correcting ordinary errors. These errors can go undetected even among careful analysts, given the large number of analytic decisions on which quantitative studies and their empirical findings are based. Because peer review can typically examine only a fraction of these decisions (those

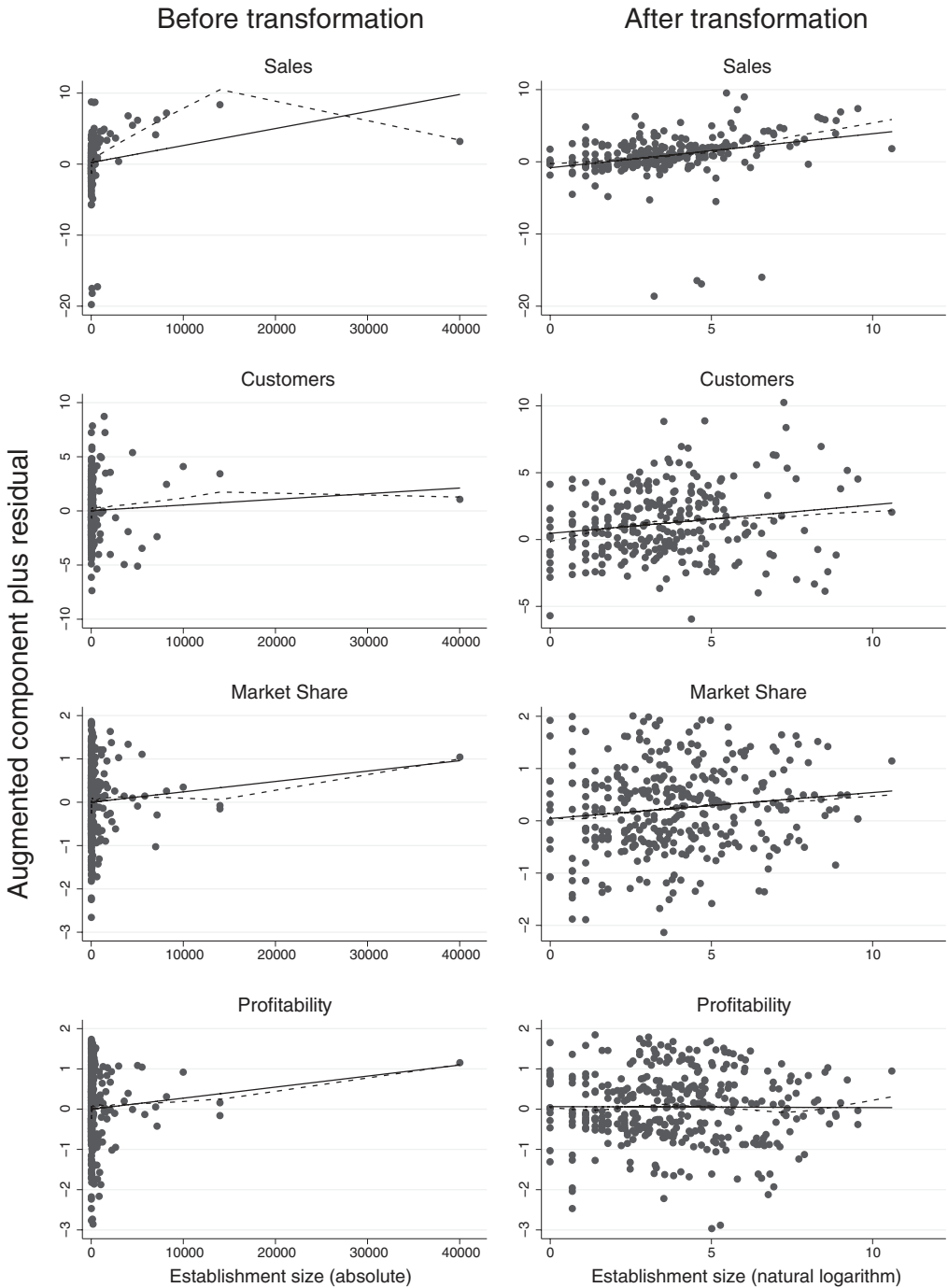


Figure 2. Component Plus Residual Plots for Establishment Size and the Four Business Outcomes

described in the paper), replication constitutes a complementary mechanism for scrutinizing and, if necessary, correcting empirical

findings. This is particularly important for findings that are as rife with policy implications as Herring’s.

Table 3. Effects of Racial and Gender Diversity on Sales, Number of Customers, Market Share, and Profitability

	Sales			Customers			Market Share			Profitability		
	Herring	SBL 1	SBL 2	Herring	SBL 1	SBL 2	Herring	SBL 1	SBL 2	Herring	SBL 1	SBL 2
Racial Diversity	.093** (3.88)	.101** (3.88)	.028 (1.08)	433.86**	8,192.5 (.77)	.027 (1.31)	.007*	.0110* (2.18)	.003 (.47)	.006+	.010+ (1.91)	.004 (.67)
Gender Diversity	.028*	.002 (.10)	-.015 (-1.03)	195.642*	6,170.6 (.94)	.032** (2.81)	.001	.001 (.41)	-.001 (-.32)	.005*	.005 (1.39)	.003 (.82)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	4.998** (.14)	.179 (.14)	-2.305+ (-1.95)	61,545.4	-446,676.1 (-.84)	3.848** (4.04)	3.403**	2.928** (10.68)	2.662** (9.66)	3.363**	3.326** (12.12)	3.096** (11.10)
N	506	239	239	506	270	266	469	348	348	484	362	362
R ²	.165	.193	.319	.155	.148	.446	.075	.137	.164	.064	.112	.123

Note: National Organizations Survey 1996/97, own calculations. “SBL” stands for Stojmenovska-Bol-Leopold. The results under “Herring” are obtained from Table 3 in Herring (2009:218). All control variables modeled after those described in Herring (2009:218). Racial diversity and gender diversity are measured using the Asymmetrical Index of Diversity (AID, see Herring 2009:214). Coefficients are unstandardized; *t*-statistics are in parentheses. +*p* < .10; **p* < .05; ***p* < .01 (two-tailed tests).

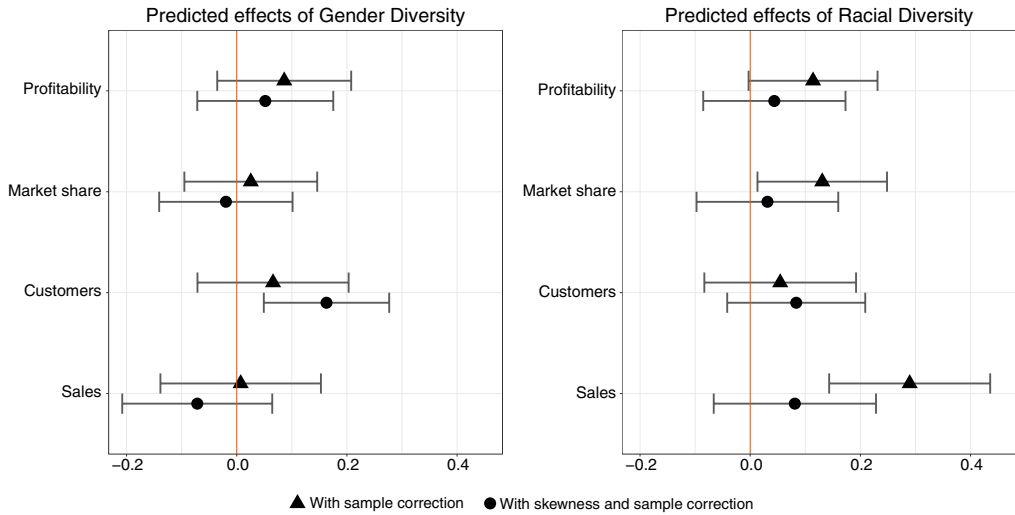


Figure 3. Predicted Effects of Gender Diversity and Racial Diversity on the Four Business Outcomes

Note: NOS 1996/97, own calculations. Standardized effects and 95% confidence intervals are shown.

Previous replications have indicated that ordinary errors occur frequently (e.g., Hernon, Ash, and Pollin 2014; McCrary 2002; for a recent example in sociology, see Karraker and Latham 2015). Several authors have called for replication standards in the social sciences that are designed to discover these errors early in the research process, but also after an article has been published (Firebaugh 2007; Freese 2007). Although our discipline has been slow to answer this call (Young 2015), some authors and journals are now aiding replication by documenting the research process more thoroughly and publishing replication data along with the articles (Zenk-Möltgen and Lepthien 2014).

Apart from the role of authors and journals, the present case illustrates how a third group of stakeholders—data producers—can contribute to the quality and reproducibility of social science research. In the NOS data analyzed by Herring, the coding conventions used for missing values are conducive to errors, given that a combination of negative and high positive values (some of which fell within the possible range of substantive codes) were used. Similar coding conventions can still be found in recent releases of large-scale survey data, such as PIAAC (OECD

2013). Positive examples of data producers who help avoid such errors include SHARE (Börsch-Supan 2015) and NEPS (Blossfeld, Roßbach, and Maurice 2011), who offer Stata ado's that automatically decode all missing values.

Beyond the shared responsibility of these stakeholders in preventing and identifying errors, our comment also demonstrates the benefits of “harnessing the undiscovered resource of student research projects” (Grahe et al. 2012) to test the replicability of published findings. This comment shows how graduate courses on replication can contribute to the cumulative progress of sociological knowledge and improve the body of existing empirical evidence.

Acknowledgments

For comments and suggestions, we thank Lonneke van den Berg, Andrea Forster, Wouter Schakel, Astrid Zwinkels, seminar participants at the University of Amsterdam, and the *ASR* editors.

Data Disclaimer

The data and tabulations utilized in this comment were made available by the archive or agency that distributed the data. Data for the 1996 to 1997 National Organizations Survey were originally collected by the Minnesota

Center for Survey Research at the University of Minnesota. Neither the original sources or collectors of the data nor the distributor of the data bear any responsibility for the analyses or interpretations presented herein.

Notes

1. The replication package to this article is available with the online version of this article and at the authors' websites (<http://www.stojmenovska.com>; <http://www.thijsbol.com>; <http://www.thomasleopold.eu>). The data are available for download at <http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/3190>.
2. This underestimation resulted in a Type II error (failing to detect an effect that is present) for company size and establishment size, and a Type I error (detecting an effect that is not present) for the diversity measures that share variance with company size and establishment size.
3. Herring develops a new measure of gender and racial diversity, the Asymmetrical Index of Diversity (AID). Many measures of diversity regard a 50/50 distribution of two groups as the maximum level of diversity. In contrast, AID does not treat underrepresentation and overrepresentation of the subordinate groups (non-whites, women) as equivalent. AID is calculated by two formulas: $AID = (1 - S)$ if $S > P$ and $AID = (1 - P)$ if $P \geq S$. S is the proportion of workers from the superordinate group (whites, men), and P is the population average of the superordinate group.
4. Establishment size is the size of the organization surveyed by the NOS. This establishment might be part of a larger company, which is then denoted by company size. If the organization in the survey is not part of a bigger company, establishment size and company size are the same.
5. The Asymmetrical Index of Diversity is constructed such that a score can range from 0 (no non-whites, no women) to the NOS 1996/97 sample average (in Herring, 25 percent for racial diversity and 46 for gender diversity). This measure is also skewed, given that all organizations that employ more non-whites or women than the sample average are set to the sample average. In additional analyses, we addressed this skew and explored other measures of diversity. These analyses are included in the replication package that accompanies our comment.
6. Similar results are found for company size (code to these plots is included in the replication package).
7. We were able to replicate about half of the effects of the control variables on the four outcomes ($4 \times 21 = 84$ effects). Thirty-five percent of the parameter estimates are in the opposite direction of those presented in Herring's article: we found 18 percent to be statistically insignificant that were significant in Herring's article, and we found 12 percent to be significant that were insignificant in Herring's article. These discrepancies are largely explained by using transformations of the skewed variables in the replication.

8. In the course, all students could select a paper of their choice. Of the five participants, four were able to replicate the substantive findings from the studies they selected.

References

- Allison, Paul D. 1999. *Multiple Regression: A Primer*. Newbury Park, CA: Sage.
- Berry, William D. 1993. *Understanding Regression Assumptions*. Newbury Park, CA: Sage.
- Blossfeld, Hans-Peter, Hans-Günther Roßbach, and Jutta Maurice. 2011. "Education as a Lifelong Process: The German National Educational Panel Study (NEPS) [Special Issue]." *Zeitschrift für Erziehungswissenschaft* 14.
- Börsch-Supan, Axel. 2015. *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5*. Release version: 1.0.0. SHARE-ERIC. Dataset, 2015. doi: 10.6103/SHARE.w5.100.
- Broockman, David, Joshua Kalla, and Peter Aronow. 2015. "Irregularities in LaCour (2014)." Unpublished paper, Stanford University, Stanford, CA.
- Firebaugh, Glenn. 2007. "Replication Data Sets and Favored-Hypothesis Bias: Comment on Jeremy Freese (2007) and Gary King (2007)." *Sociological Methods & Research* 36(2):200–209.
- Freese, Jeremy. 2007. "Replication Standards for Quantitative Social Science: Why Not Sociology?" *Sociological Methods & Research* 36(2):153–72.
- Fox, John. 1991. *Regression Diagnostics: An Introduction*. Newbury Park, CA: Sage.
- Grahe, Jon E., Alan Reifman, Anthony D. Hermann, Marie Walker, Kathryn C. Oleson, Michelle Nario-Redmond, and Richard P. Wiebe. 2012. "Harnessing the Undiscovered Resource of Student Research Projects." *Perspectives on Psychological Science* 7(6):605–607.
- Herndon, Thomas, Michael Ash, and Robert Pollin. 2014. "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff." *Cambridge Journal of Economics* 38(2):257–79.
- Herring, Cedric. 2009. "Does Diversity Pay? Race, Gender, and the Business Case for Diversity." *American Sociological Review* 74(2):208–224.
- Kalleberg, Arne L., David Knoke, and Peter V. Marsden. *National Organizations Survey (NOS), 1996–1997* [Computer file]. ICPSR version. Minneapolis: University of Minnesota Center for Survey Research [producer], 1997. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001.
- Karraker, Amelia, and Kenzie Latham. 2015. "Authors' Explanation of the Retraction." *Journal of Health and Social Behavior* 56(3):417–19.
- King, Gary. 2006. "Publication, Publication." *PS: Political Science & Politics* 39(1):119–25.
- McCrary, Justin. 2002. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment." *American Economic Review* 92(4):1236–43.

- OECD. 2013. *Programme for the International Assessment of Adult Competencies (PIAAC)*. Paris: OECD.
- Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349(6251).
- Osborne, Jason W., William R. Christianson, and Jason S. Gunter. 2001. "Educational Psychology from a Statistician's Perspective: A Review of the Power and Goodness of Educational Psychology Research." Presented at the national meeting of the American Education Research Association (AERA), Seattle, WA.
- Young, Cristobal. 2015. "Do Sociologists Release Their Data and Code? Disappointing Results from a Field Experiment on Replication." *Orgtheory.net* Blog, August 11. Retrieved April 29, 2016 (<https://orgtheory.wordpress.com/2015/08/11/sociologists-need-to-be-better-at-replication-a-guest-post-by-cristobal-young/>).
- Zenk-Möltgen, Wolfgang, and Greta Lepthien. 2014. "Data Sharing in Sociology Journals." *Online Information Review* 38(6):709–722.

Dragana Stojmenovska is a PhD Candidate at the University of Amsterdam, affiliated with the Amsterdam Institute for Social Science Research and the Amsterdam Research Center for Gender and Sexuality. Her research affinities lie with gender, occupational segregation, and gendered social structures. In her PhD project, funded by

a Research Talent grant awarded by the Netherlands Organization for Scientific Research, she investigates women's underrepresentation in organizational authority from an organizational perspective.

Thijs Bol is Assistant Professor in the Department of Sociology at the University of Amsterdam. His research interests are in occupations, inequality, the transition from school to work and social stratification. In a current research project he investigates school-to-work linkages from a country-comparative perspective. Recent publications include "School-to-Work Linkages in the United States, Germany, and France" (*American Journal of Sociology* 2017) and "Vocational Education and the Employment over the Life-Cycle" (*Sociological Science* 2016).

Thomas Leopold is Associate Professor in the Department of Sociology at the University of Amsterdam. His research interests are in family sociology, life-course research, demography, and social stratification. Recent publications include "The Demography of Grandparenthood: An International Profile" (*Social Forces* 2015), "Convergence or Continuity? The Gender Gap in Household Labor after Retirement" (*Journal of Marriage and Family* 2015), and "Is Divorce More Painful When Couples Have Children? Evidence from Long-Term Panel Data on Multiple Domains of Well-Being" (*Demography* 2016).