



Københavns Universitet



Regression to Causality

Bordacconi, Mats Joe; Larsen, Martin Vinæs

Published in:
Research and Politics

DOI:
[10.1177/2053168014548092](https://doi.org/10.1177/2053168014548092)

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Bordacconi, M. J., & Larsen, M. V. (2014). Regression to Causality: Regression-style presentation influences causal attribution. DOI: 10.1177/2053168014548092

Research & Politics

<http://rap.sagepub.com/>

Regression to causality: Regression-style presentation influences causal attribution

Mats Joe Bordacconi and Martin Vinæs Larsen

Research & Politics 2014 1:

DOI: 10.1177/2053168014548092

The online version of this article can be found at:
<http://rap.sagepub.com/content/1/2/2053168014548092>

Published by:



<http://www.sagepublications.com>

Additional services and information for *Research & Politics* can be found at:

Email Alerts: <http://rap.sagepub.com/cgi/alerts>

Subscriptions: <http://rap.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Sep 1, 2014

[What is This?](#)

Regression to causality: Regression-style presentation influences causal attribution

Research and Politics
 July-September 2014: 1–6
 © The Author(s) 2014
 DOI: 10.1177/2053168014548092
 rap.sagepub.com


Mats Joe Bordacconi and Martin Vinæs Larsen

Abstract

Humans are fundamentally primed for making causal attributions based on correlations. This implies that researchers must be careful to present their results in a manner that inhibits unwarranted causal attribution. In this paper, we present the results of an experiment that suggests regression models – one of the primary vehicles for analyzing statistical results in political science – encourage causal interpretation. Specifically, we demonstrate that presenting observational results in a regression model, rather than as a simple comparison of means, makes causal interpretation of the results more likely. Our experiment drew on a sample of 235 university students from three different social science degree programs (political science, sociology and economics), all of whom had received substantial training in statistics. The subjects were asked to compare and evaluate the validity of equivalent results presented as either regression models or as a *t*-test of two sample means. Our experiment shows that the subjects who were presented with results as estimates from a regression model were more inclined to interpret these results causally. Our experiment implies that scholars using regression models should note carefully both their models' identifying assumptions and which causal attributions can safely be concluded from their analysis.

Keywords

Attribution, cognitive biases, political methodology

Introduction

'Regression models make it all too easy to substitute technique for work.' (Freedman, 1991: 300)

Everybody – probably – knows that correlation does not imply causation. Even so, we are often misled by a captivating correlation. This is not surprising. Individuals have a fundamental cognitive need for making causal attributions (Heider, 1958). However, when it comes to conducting research, we often go to great lengths to distinguish when we are referring to causation from when we are simply referring to correlation. This is especially true when it comes to observational data, where dangers such as confounders, classical measurement error and reverse causality threaten the prospects of causal identification at every turn (Winship and Morgan, 1999). Even so, results of observational studies in political science and other disciplines are often presented in a manner which seems to be skewed towards causal inference: regression models.

Regression models are characterized by a parameterization of the relationship between an outcome variable and a set of regressor variables. The asymmetric nature of the regression model, where a 'dependent' variable is set as a

function of a set of 'independent' variables, can lead one to think in terms of causal inference, thus implying that the regressors are predetermined *and* that they determine the outcome variable. The standard way of presenting the results of a regression model heightens this perception: the estimated parameter(s) of each regressor is/are presented next to the variable itself and presented as the *ceteris paribus* effect of this regressor on the outcome variable. Taken together, this heavy semantic skew towards causal inference might very well skew readers towards making causal inferences, even if the results are rife with concerns of endogeneity.

We investigate below whether regression models have such a *causal skew*, by looking at how the use of regression models influence subjects' interpretation of statistical results from observational studies.

Department of Political Science, University of Copenhagen, Denmark

Corresponding author:

Martin Vinæs Larsen, Department of Political Science, University of Copenhagen, Øster Farimagsgade 5, Copenhagen, DK-135, Denmark.
 Email: mvl@ifs.ku.dk



Creative Commons Non Commercial CC-BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 License (<http://www.creativecommons.org/licenses/by-nc/3.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<http://www.uk.sagepub.com/aboutus/openaccess.htm>).

The main study: The effect of regression-style presentation

To test whether regression-style presentation of results leads one to interpret these results in a more causal fashion, we conducted an experiment in which subjects were shown the results of two different (fabricated) studies. Although the results were the same for all respondents, we randomized whether the analysis was presented as a linear regression of a dummy variable on a continuous variable or as an independent sample *t*-test of the mean difference between the sample averages of two groups. The two statistical techniques are equivalent, because linear regression reduces to a comparison of sample means if the only regressor is a single dichotomous variable.¹ As such, the difference in interpretation cannot stem from greater confidence in one set of result over another, and instead must stem from how the results are presented. We describe below the experiment's subjects, questionnaire and conditions, and then we turn to the results.

Subjects

235 social science students at the University of Copenhagen participated in the experiment. The students were in their second or third year of the degree program and had received extensive training in the presented statistical techniques: 83 were recruited from the political science program, 83 from economics and 63 from sociology. The experiment was conducted in a class room setting, and it consisted of filling out a short questionnaire while we were present. However, the authors were not affiliated with the students in any other way. We chose students as our study participants because they were familiar with the statistical techniques presented in the questionnaire; because it is intrinsically interesting to understand how students consume scientific work; and because students are probably the largest block of consumers when it comes to scientific work (*vis-à-vis* other researchers or professionals).

Questionnaire and conditions

The questionnaire initially presented the results of a seemingly real study. Specifically, subjects were told that 'a survey of 1053 representative Danish voters had examined the relation between how much you like the Social Democrats and ones news consumption. See table 1 for details.' The subjects were then presented with one of two tables in figure 1 at random. Both tables presented identical results, either as a regression model or as an independent sample *t*-test. Afterwards, subjects were asked to evaluate whether they would accept the following two statements, based on the presented results.

1. *News consumption affects sympathy for the Social Democrats.*

Table 1. Descriptive statistics.

T-test	Regression		
	M	SD	n
News – Sig.	3.97	1.90	113
News – Caus.	3.48	1.92	112
Dem. – Sig.	2.20	1.52	112
Dem. – Caus.	4.73	1.98	106

2. *It is likely that the differences in sympathy for the Social Democrats are due to chance.*

Two additional statements were posed in a similar manner to disguise the purpose of the study. Responses to these questions were recorded on a scale from one ('Would not accept the statement at all') to seven ('Would completely accept the statement'). The key dependent variable in this analysis is drawn from the responses to the first statement, as it implies a causal relationship between the two variables. It is what we will use to gauge respondents' causal beliefs. The second statement simply implies a statistically insignificant association, and we use these responses for a placebo test below.

The subjects then did the same for a second study. In Study 2, the identified relationship was between democracy (measured dichotomously) and GDP per capita (measured continuously) (see the Appendix for the exact wording of all the questions). Finally, subjects were asked about their age and gender; the average age was 23, and 51% were male). Taken together each group of subjects spent about ten minutes on the survey. As mentioned above, both studies were fabricated for use in this study. Subjects were informed of this after completing the questionnaire.

The treatment effect

To identify the effect of our experimental conditions on causal interpretation, we model each subject's belief in the causality and significance statements as a linear function of: (1) whether the subject received the results in a regression-style format; and (2) in which degree program they were enrolled (economics, political science or sociology). The key parameter to be estimated is the effect of receiving the regression-style presentation, which signifies the mean difference between those who received the results in a regression-style format and those who received the results in a *t*-test format. Figure 2 presents this estimate, which was obtained using an OLS regression. Table 1 presents additional information on the dependent variables across experimental conditions.

Figure 2 reveals a statistically significant and positive effect of regression-style presentation on causal interpretation across both studies presented in the questionnaire: democracy and news consumption. Subjects receiving results

Table 1: OLS regression	
	Sympathy for the Social Democrats
Watch the news several times a week (Reference: Doesn't watch the news)	-0.2* (0.1)
Constant	0.7***
n	1053

Notes: Standard error in parenthesis.
Significance level: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.
Question: How much do you like the Social Democrats? (1=Very much, 0=Not at all)

Table 1: Independent sample t-test	
	Sympathy for the Social Democrats
Watch the news several times a week	0.5
Doesn't watch the news	0.7
Average difference	-0.2* (0.1)
n	1053

Notes: Standard error in parenthesis.
Average differences tested by using independent sample t-test. Assumption about equal variance not violated.
Significance level: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.
Question: How much do you like the Social Democrats? (1=Very much, 0=Not at all)

Figure 1. Experimental treatments.

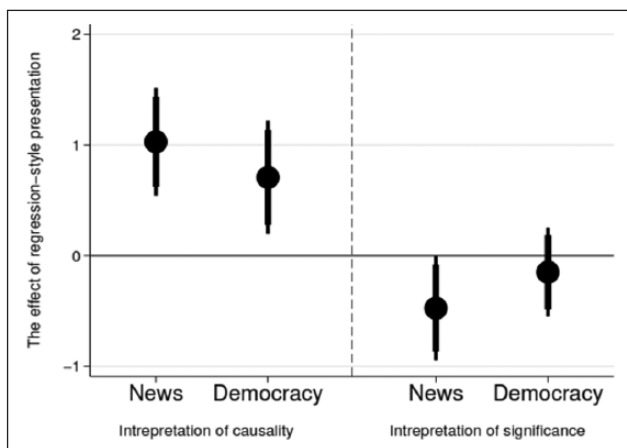


Figure 2. The effect of regression-style presentation.
Note: Thick spikes signify the 90% Confidence Intervals (CIs) and thin spikes signify the 95% Confidence Intervals.

in regression-form were more likely to deem the relationship found in the studies causal. For both studies the estimate is close to one, suggesting that the manipulation moved them one point up (or down) on the seven-point scale.

The same is not true for beliefs about a statistically significant relationship. While subjects found it somewhat less likely that the relationship between variables was statistically insignificant when results were presented in regression form, in the study on news consumption the effect was

not large enough to achieve statistical significance at the 5% level, while in the study on democracy the effect was small to achieve statistical significance at any acceptable level ($p > 0.4$).

Heterogeneity of the treatment effect

On average, regression-style presentation led subjects to interpret results more causally, but the question is raised: were certain types of students more affected by the regression-style presentation than others? This is potentially interesting, because it can tell us something about the kinds of individuals to whom this sort of biased information processing applies.

In order to investigate such heterogeneity, we estimated the difference in the average marginal effects across programs and age. Figure 3 presents these results for both the democracy and the news consumption case. Only one group of students had markedly different treatment effects: older sociology students. For this group the effect actually goes in the opposite direction, although the estimate is somewhat imprecise.

This result cannot be explained by the level of statistical training. The economics students in our sample had, by far, gotten the most training in statistics, with two courses on math and statistics and two courses on econometrics, but they did not distinguish correlation from causation at higher rates than other students. One potential

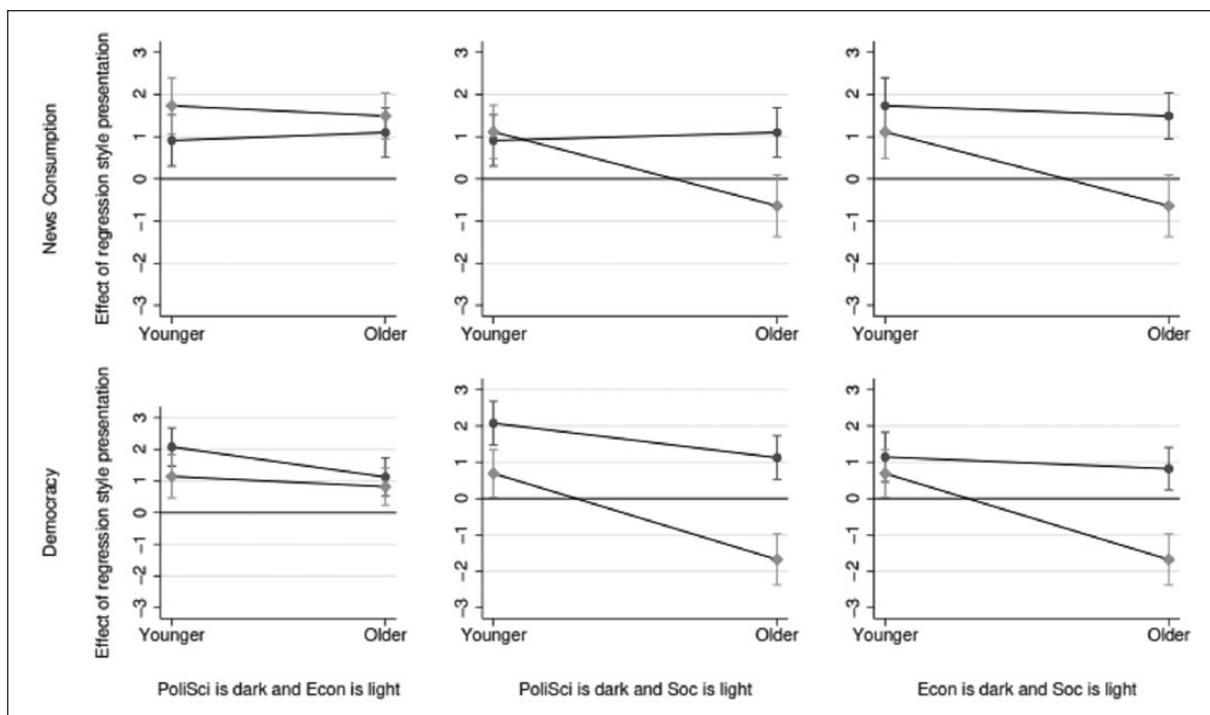


Figure 3. Heterogeneity of the treatment effect.

Note: Capped spikes signify 68% CIs of the effect, representing one standard error. Effects and their variance were estimated using OLS with a three-way interaction of treatment, program and age. 'Older students' are those of 23 years of age or above.

explanation might be that the sociology students who participated in our experiment had more recently finished the statistics course in which they first heard of OLS regression and *t*-test. As such, they had recently been 'reminded' that there is no difference between the two techniques, which might have mitigated the bias in line with predictions by Kahneman (2003: 1454–1456). This explanation, however, does not account for older sociology students experiencing different treatment effects than their younger counterparts.

In summary, then, we see little heterogeneity in the treatment effect across age and program, and the differences we do see are unfortunately not easy to interpret.

Scale validation

The experiment we present above reveals that regression-style presentation of results led people to interpret the results more causally. Critics of our experiment might argue that the causal statement, which was meant to gauge respondents' belief that the results presented to them stemmed from a causal relationship, did not actually measure any such belief. These critics might argue that the statement instead simply measured subjects' belief in a correlation, in the credibility of the researcher, or in something completely different. To investigate this, we conducted a small follow-up study with a comparable set of respondents.

Table 2. Does the statement imply causal effect

	News consumption		Democracy	
	M	SE	M	SE
Causal statement	7.7	0.5	8.2	0.3
Significance statement	1.7	0.4	2.0	0.5

Specifically, 26 political science students, who did not participate in the main study, filled out a short questionnaire. The participants had received roughly the same statistical training as the political science students in the main study, but were from another cohort. The questionnaire was filled out in a class-room setting. It presented the same causal and significance statements used in the first experiment. For each statement, respondents were asked whether the statement implied a causal effect of the independent variable in question (i.e. news consumption or democracy) on the dependent variable (i.e. party sympathy and GDP per capita). Responses were recorded on a scale from zero ('Does not imply a causal effect at all') to 10 ('Strongly implies a causal effect'). The questionnaire explicitly instructed respondents to use the conception of causality they had learned in their methodology and statistics classes.

The results are presented in Table 2. Respondents' mean response to the causal statements was about 8 (out of 10),

whereas their mean response to the significance statements was about 2.

This follow-up experiment allows us to conclude that when the subjects in the main study reacted to the regression style presentation by agreeing to the causal statement, they agreed to a statement which implies a causal effect in a statistical sense.

Discussion

That frames sometimes matter just as much as what they are framing is the foundation of much research in cognitive psychology, ranging from effects of social comparisons (Mussweiler, 2003), to anchors (Strack and Mussweiler, 1997) and priming effects (Higgins and Brendl, 1995). In this experiment we documented a new way in which frames affect what they are framing: namely, how the presentation of statistical results influences the way readers interpret these results. Specifically, a regression-style presentation lead statistically savvy subjects to interpret results more causally than results presented as independent sample *t*-tests. In other words, regression models – one of the main ways of presenting the findings of observational studies – bias readers towards believing that the identified associations are causal. Our experiment was less conclusive as to which part of the regression-style presentation actually made a difference in how subjects interpreted findings.

We offer two potential explanations for these findings: first, that the language of effects used when reading a regression table primes considerations of causality; and, second, that the more technical nature of the regression analysis, compared to just testing differences in means, lends more legitimacy to the presented claims. The second explanation would be in line with recent studies by Eriksson (2012) and Weisberg (2008), who suggest that as explanations become more technical and complex, they are perceived as more believable.

Other explanations are viable but do not seem as likely. For example, one might suggest that the researcher signals a belief in a specific causal direction when analyzing the results with a regression, as one must specify a regressor and outcome variable in regression analysis but need not do so when comparing means. Accordingly, respondents' reactions might simply reflect different signals about the researcher's beliefs. While this is certainly possible, we doubt that it explains our findings, because a researcher can also designate an independent variable by picking a grouping variable to be the basis of the comparison in an independent sample *t*-test.

Finally, one might argue that the findings are a result of simple statistical illiteracy: the subjects actually believed that an OLS regression is better at delineating causation. We do not find this latter explanation to be very convincing. All subjects in our study had received intense training in the

presented techniques, and this training was explicitly focused on causal inference. Furthermore, the economics students, who had received the most training in statistics, were just as prone to the bias of inferring causation from regression as the political science students, and were actually even more prone to this bias than some sociology students. Further research must work to delineate the exact causal mechanisms at work.

Our findings do not suggest that we should abandon regression models as a way of analyzing our data. Indeed, it is hard to imagine observational studies without regression modelling. Our experiment does, however, imply that scholars using regression models should carefully note both their models' identifying assumptions and which causal attributions can safely be concluded from their analysis. A more broad implication of our findings is that scholars should present results in several ways. Presenting predicted probabilities and graphical representations alongside regression models would allow scholars to paint a more nuanced picture of their results.

Acknowledgement

The authors would like to thank the anonymous reviewers and Jens Olav Dahlgaard, Jakob Gerner Hariri and, especially, Asmus Leth Olsen for valuable comments on the note.

Declaration of conflicting of interest

The authors declare that there are no conflicts of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Note

1. While presenting bivariate analyses, such as 'simple' linear regression, potentially limit the scope of our experiment, it is necessary to make the two kinds of treatments comparable. Further, one should keep in mind that results from multiple regressions are often preceded by bivariate analyses.

References

- Eriksson Kimmo (2012) The nonsense math effect. *Judgment and Decision Making* 7(6): 746–749.
- Freedman DA (1991) Statistical models and shoe leather. *Sociological Methodology* 21(2): 291–313.
- Heider F (1958) *The Psychology of Interpersonal Relations*. New York, NY: John Wiley.
- Higgins ET and Brendl CM (1995) Accessibility and applicability: Some 'activation rules' influencing judgment. *Journal of Experimental Social Psychology* 31(3): 218–243.
- Kahneman D (2003) Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review* 93(5): 1449–1475.

- Mussweiler Thomas (2003) Comparison processes in social judgment: mechanisms and consequences. *Psychological Review* 110(3): 472.
- Strack F and Mussweiler T (1997) Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology* 73(3): 437.
- Weisberg DS, Keil FC, Goodstein J, Rawson E and Gray JR (2008) The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience* 20(3): 470–477.
- Winship C and Morgan SL (1999) The estimation of causal effects from observational data. *Annual Review of Sociology* 25: 659–706.

Appendix: Additional information on the survey

The four statements made about Study 1 were:

1. *News consumption affects sympathy for the Social Democrats;*
2. *It is likely that the differences in sympathy for the Social Democrats are due to chance;*
3. *All else being equal, people who do not watch news several times a week have an average sympathy for the Social Democrats of 0.7; and*
4. *People who watch the news several times a week like the Social Democrats better than those who do not.*

The four statements made about Study 2 were

1. *Democracy affects a country's GDP per capita;*
2. *It is likely that the difference in GDP per capita is due to chance;*
3. *All else being equal, autocratic countries have an average GDP per capita of US\$20,000; and*
4. *The citizens of autocratic countries are poorer than the citizens of democratic countries.*

Figure 4. Experimental treatments, Study 2.

Table 2: OLS regression	
	GDP per capita
Democracy (Reference: Autocracy)	20.000*** (3.500)
Constant	20.000***
n	220

Notes: Standard error in parenthesis.
Significance level: * = p<0.05, ** = p<0.01, *** = p<0.001.

Table 2: Independent sample t-test	
	GDP per capita
Democracy	40.000
Autocracy	20.000
Average difference	20.000*** (3.500)
n	220

Notes: Standard error in parenthesis.
Average differences tested by using independent sample t-test. Assumption about equal variance not violated.
Significance level: * = p<0.05, ** = p<0.01, *** = p<0.001.