

Response to Powell and Tucker’s “Little is Known about Party System Volatility in Post-Communist Europe, but We have Interesting New Methods and Data for Studying It”

CHARLES CRABTREE
University of Michigan

MATT GOLDER
Pennsylvania State University

We appreciate Powell and Tucker’s response to our research note (Crabtree and Golder, 2016) reexamining their 2014 empirical analysis of the determinants of party system volatility in Eastern Europe (Powell and Tucker, 2014). We now take this opportunity to respond to several issues they raise.

Similar Substantive Claims?

Powell and Tucker (2016, 1) state that we provide “further evidence in support of the primary substantive conclusion of [their] article.” However, this is not accurate. Specifically, our research note “challenge[s] their central claim that replacement volatility in post-communist Europe is driven by long-term economic performance” (Crabtree and Golder, 2016, 1). In their response, Powell and Tucker (2016, 2) downplay the importance of this particular claim, arguing that their primary substantive conclusion was that “little is known about the correlates of either [replacement or electoral] volatility in the first two decades of post-communist elections.” We believe their argument here is at odds with what they say in their original article.

In their original article, Powell and Tucker (2014) disaggregate party system volatility into two components: (i) ‘replacement volatility’ caused by new party entry and old party exit, and (ii) ‘electoral volatility’ caused by vote switching across existing parties.¹ Whereas Powell and Tucker (2014, 123) conclude that scholars know little about the determinants of *electoral volatility* in post-communist Europe, they explicitly claim that *replacement volatility* is “largely a function of long-term economic recovery.” This particular

¹Powell and Tucker (2014) refer to these two types of volatility as “Type A Volatility” and “Type B Volatility.” These names are not particularly informative, and so we prefer, in line with Birch (2001, 2003), to refer to them as replacement volatility and electoral volatility.

claim appears in the article’s abstract, and is, in fact, the only substantive conclusion reported in the abstract.

The abstract states that,

“entry and exit [replacement] volatility is found to be largely a function of long-term economic recovery, and it becomes clear that very little is known about what causes ‘party switching’ [electoral] volatility” (Powell and Tucker, 2014, 123).

As this indicates, it is not the case that Powell and Tucker (2014) conclude that little is known about what causes *both* types of volatility — they conclude that replacement volatility is driven by long-term economic recovery.

In fact, Powell and Tucker (2014) consider their claim that replacement volatility is driven by long-term economic recovery to be sufficiently important that it appears as one of the four ‘walk-away’ conclusions they highlight at the end of their article. They write that they have,

“four important conclusions concerning . . . volatility in post-communist countries. . . our third conclusion is that, unlike previous studies of . . . volatility in post-communist countries, we find almost no support for the claim that structural factors such as district magnitude, electoral rules or ethnic heterogeneity affect levels of . . . volatility across elections, and only marginal evidence in support of the governing system. Nor does party system fractionalization seem to matter. Instead, we find just one statistically significant predictor of [replacement] volatility over the first two decades of post-communist elections: the worse the economy is performing relative to where it was at the start of the transition, the more likely [replacement] volatility is to be high” (Powell and Tucker, 2014, 142-143).

Moreover, they go on to argue that this particular empirical finding has a strong basis in theory. Specifically, they note that it,

“fits in very well with big-picture theoretical ideas about the extended role of the transition in structuring post-communist political behavior (Kitschelt et al., 1999; Tucker, 2006): a country’s long-term economic performance seems to create space for the creation of new political parties and incentives for failed ones to be abandoned” (143).

As these quotes indicate, it is a central substantive conclusion from Powell and Tucker’s (2014) empirical analysis that replacement volatility is driven by long-term economic recovery.² Our research note

²Contrary to the claims of Powell and Tucker (2016, 4), our research note does not claim that the empirical results from Powell and Tucker’s analysis of the determinants of party system volatility represent the “primary contribution” of their article. Indeed, our research note starts by recognizing that Powell and Tucker (2014) “highlight the importance of disaggregating volatility into two components” and that they show “that volatility is primarily driven by replacement volatility in the post-communist countries of Eastern Europe but by electoral volatility in the established democracies of Western Europe.” Powell and Tucker (2014) do go on, however, to examine the determinants of the two different types of volatility in post-communist Europe. It is the results of this empirical analysis that we examine in our research note.

challenges this conclusion as it finds no statistically significant evidence that long-term economic recovery influences replacement volatility. Far from providing evidence in support of their primary substantive conclusion, our analysis directly challenges it.

[Powell and Tucker \(2014, 4\)](#) seek to downplay the importance of our empirical challenge by arguing that “only six (8 per cent) [of the 83 articles that had cited their original piece by April 2015] referenced the finding that economic conditions impacted . . . volatility.” We find this particular reasoning somewhat troubling. Another way to look at these numbers is that within a year of Powell and Tucker publishing their paper, six articles had already been written that drew on a result that we now know lacks empirical evidence and was based on miscalculated data. By June 2016, Powell and Tucker’s original article had received a further 57 citations. At the rate calculated by [Powell and Tucker \(2016\)](#), this suggests that a further five articles are now in the same boat. We think this is problematic and worthy of correction.

Statistical Significance?

The core finding in our research note is that Powell and Tucker are incorrect to claim that replacement volatility in post-communist Europe is driven by long-term economic performance. This is because there is no statistically significant relationship between long-term economic performance and replacement volatility once we use the correctly calculated data for the three Bosnia observations (or simply remove the three outlier Bosnia observations). In their response, [Powell and Tucker \(2016\)](#) takes us to task for focusing solely on whether there is a *statistically* significant relationship between long-term economic performance and replacement volatility. They write that “many methodologists more sophisticated than us have long argued that too much attention is paid to standard significance levels at the expense of substantive effects, and this may serve as an example” ([Powell and Tucker, 2016, 2](#)).

We focused on statistical significance because [Powell and Tucker \(2014, 134\)](#) explicitly state that the goal of their empirical analysis was “to see if we can recover a set of statistically significant predictors of . . . volatility.” In effect, we chose to use the same criterion that Powell and Tucker adopted to evaluate their own empirical results. In line with their stated goal, [Powell and Tucker \(2014\)](#) focus on statistical significance when describing their results. Indeed, they have seventeen distinct sentences that reference the statistical significance of their results; none discuss substantive significance. Our research note simply indicates that Powell and Tucker’s central claim about the driving force behind replacement volatility in

post-communist Europe is not supported by their own chosen criterion of evaluation.

Powell and Tucker are, of course, correct that substantive significance matters, and this is a point that we will return to shortly. However, this does not mean that statistical significance is irrelevant. The American Statistical Association recently published an unprecedented statement on statistical significance and p -values (ASA, 2016). The goal of the statement was to “shed light on an aspect of the field that is too often misunderstood and misused in the broader research community” (Wasserstein and Lazar, 2016). The statement proposes six “principles.” The very first principle indicates that statistical significance and p -values “can be a useful statistical measure” (ASA, 2016, 131). It states that,

“ P -values can indicate how incompatible the data are with a specified statistical model.

A p -value provides one approach to summarizing the incompatibility between a particular set of data and a proposed model for the data. The most common context is a model, constructed under a set of assumptions, together with a so-called ‘null hypothesis.’ Often the null hypothesis postulates the absence of an effect, such as no difference between two groups, or the absence of a relationship between a factor and an outcome. The smaller the p -value, the greater the statistical incompatibility of the data with the null hypothesis, if the underlying assumptions used to calculate the p -value hold. This incompatibility can be interpreted as casting doubt on or providing evidence against the null hypothesis or the underlying assumptions” (ASA, 2016, 131).

The p -value associated with the coefficient on *GDP Change from 1989* (the variable measuring long-term economic performance) reported by Powell and Tucker (2014) is $p = 0.002$, suggesting that their original data are inconsistent with the null hypothesis of no relationship between long-term economic performance and replacement volatility. This is the basis for their claim that replacement volatility is driven by long-term economic performance. However, the p -value associated with this coefficient jumps considerably to $p = 0.38$ once we correct the miscalculated values on *GDP Change from 1989* for the three Bosnia observations. The p -value is even larger ($p = 0.41$) if we simply drop the three outlier Bosnia observations. Note that these p -values are considerably larger than the standard significance levels to which Powell and Tucker (2016, 2) refer when they complain that “too much attention is paid to standard significance levels.” These new p -values indicate that Powell and Tucker’s data are consistent with the null hypothesis that there is no relationship between long-term economic performance and replacement volatility. Far from being irrelevant, this is meaningful evidence and should not be ignored.

Powell and Tucker are inconsistent when applying their criteria for evaluating empirical results. In their original analysis, they rule out all but one variable (*GDP Change from 1989*) as a determinant of

replacement volatility on the grounds that their coefficients do not achieve “statistical significance at conventional levels” (Powell and Tucker, 2014, 140). Indeed, the lack of statistically significant coefficients is precisely why Powell and Tucker (2016, 1) claim in their response that “little is known about party system volatility in post-communist Europe.” When we show in our research note, though, that the coefficient on *GDP Change from 1989* is also statistically insignificant once the three observations for Bosnia are calculated correctly (or dropped), they decline to rule out long-term economic performance as a key determinant of replacement volatility. Instead, they offer the possibility that long-term economic performance might nonetheless be substantively significant. The inconsistency arises because they never entertain this possibility for any of the other statistically insignificant results.

Finally, before turning to the substantive significance of their results, we want to respond to the new analyses that Powell and Tucker (2016, 3) present in Table 1 of their response. In these analyses, Powell and Tucker (2016) use their *original* data and sequentially drop one country at a time from a series of regressions. They report that the coefficients on *GDP Change from 1989* are always statistically significant except when Bosnia is dropped from the sample.³ They argue that this supports their original claim because it “suggests that the relationship between GDP-89 and [replacement] volatility was not so weak that dropping countries at random could produce similar effects” (Powell and Tucker, 2016, 2).

However, this misses the point. The values on *GDP Change from 1989* for the three Bosnia observations are distinct outliers and are not calculated correctly in the original data. *GDP Change from 1989* is supposed to be operationalized as the ratio of GDP in a given election year to GDP in 1989. Although they do not report it in their original manuscript or in the codebook that accompanies their replication data, Powell and Tucker (2014) actually use 1994 as the reference year for the Bosnia observations. They do this not for theoretical reasons but because they were unable to locate data on Bosnia’s GDP prior to 1994. As our research note points out, choosing 1994 as the reference year for Bosnia has significant consequences:

“The 1992-1995 Bosnian war essentially destroyed the nation’s economy, something that is clearly visible in Figure 2. Bosnia-Herzegovina’s GDP in 1994 was just 12.7% of what it was in 1989. Although GDP was more than five to six times larger in the 2000s than it was in 1994 (the values calculated by Powell and Tucker), it actually remained less than it was in 1989” (Crabtree and Golder, 2016, 5).

³It turns out that these analyses are not new — they are exactly the same as the ones we already report in our research note. We reported that, “A jackknife procedure, where we sequentially drop each country one at a time and then re-estimate the models for each of the reduced datasets, reveals that it is only the outlier observations from Bosnia-Herzegovina that are substantively influencing the results” (Crabtree and Golder, 2016, 4).

The decision to use 1994 instead of 1989 as the reference year for Bosnia results in the values on *GDP Change from 1989* for the three Bosnian observations being distinct outliers.

“Whereas the mean value of [*GDP Change from 1989*] is 5.66 for Bosnia-Herzegovina, it is only 0.98 ($\sigma = 0.39$) for the remaining 86 observations. In fact, the average value for Bosnia-Herzegovina is fully 2.78 times larger than the next largest value, Poland 2007” (Crabtree and Golder, 2016, 3-4).

As our research note demonstrates, the coefficient on *GDP Change from 1989* is not statistically significant ($p = 0.38$) once we use the correctly calculated data from Bosnia. Even if one were to use Powell and Tucker’s original data, a simple bivariate plot like the one shown in Figure 1 of our research note indicates that the three Bosnia observations are distinct outliers and should not be pooled with the other observations (Crabtree and Golder, 2016, 4). What the results in Table 1 of Powell and Tucker’s response actually indicate is that the coefficient on *GDP Change from 1989* is only statistically significant if the sample includes miscalculated and outlier data from Bosnia. We do not believe that this is a compelling empirical basis for Powell and Tucker’s (2014, 123) claim that replacement volatility is “largely a function of long-term economic recovery.”

Substantive Significance?

While the 2016 statement from the American Statistical Association recognizes that discussions of statistical significance and p -values are useful, it also indicates that they should be complemented by other analyses that focus on substantive significance. As we have already indicated, we focused solely on statistical significance in our research note because statistical significance was the specific and only criterion that Powell and Tucker (2014) chose for evaluating their empirical results. However, we now take this opportunity to demonstrate that in addition to not having a statistically significant effect on replacement volatility in post-communist Europe, long-term economic performance does not have a *substantively significant* effect either.

In their response, Powell and Tucker (2016) do not back down from their claim that replacement volatility is driven by long-term economic performance. Specifically, they write that,

“caution is still in order in claiming that the book is definitely closed on the possible link between growth since the start of the post-communist period and [replacement] volatility. . . . it remains the case that both using the new Bosnia-Herzegovina data and list-wise deleting the Bosnia-Herzegovina cases from the analysis increases the size of the coefficient on GDP-89. Many methodologists more sophisticated than we have long argued that too much attention is

paid to standard significance levels at the expense of substantive effects, and this may serve as an example” (Powell and Tucker, 2016, 2).

Although Powell and Tucker (2016) raise the possibility that long-term economic performance may have a substantively significant effect on replacement volatility, they do not test to see whether this is the case. Instead, they point to the fact that the magnitude of the coefficient on *GDP Change from 1989* increases when we use the correct data for Bosnia or if we drop the incorrect Bosnia observations.

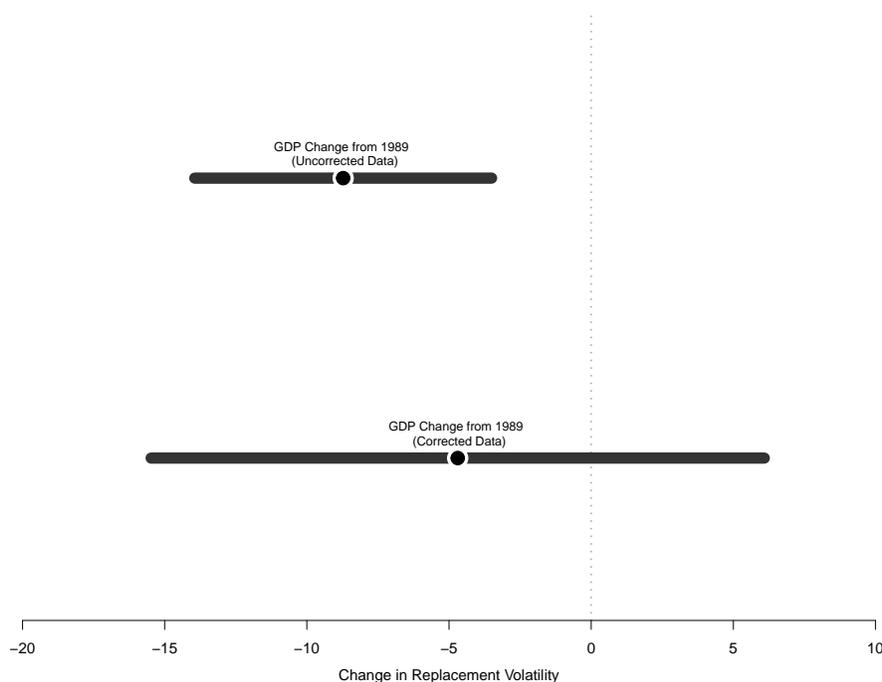
Substantive significance cannot be evaluated by simply looking at the magnitude of a coefficient in isolation, though. Among other things, one also needs to take account of the scale on which a variable is measured and the variation in the observed values of a variable. Once you do this, it becomes clear not only that the substantive effect of long-term economic performance is considerably *lower* with the corrected data than with the original data (despite the coefficient on *GDP Change from 1989* being larger with the corrected data), but also that the absolute size of any substantive effect is small.

In their original analysis, Powell and Tucker (2014) report that the coefficient on *GDP Change for 1989* is -4.62 [-7.39, -1.86], with two-tailed 95% confidence intervals shown in parentheses. The p -value associated with this coefficient is $p = 0.002$. The mean value of *GDP Change for 1989* in the original data is 1.14 with a standard deviation of 0.94; the minimum and maximum observed values are 0.22 and 6.66. Given the magnitude of the estimated coefficient on *GDP Change for 1989*, the substantive effect of moving from one standard deviation below the mean value on *GDP Change for 1989* to one standard deviation above (a swing of two standard deviations) is to reduce replacement volatility by 8.72 [3.50, 13.94].

Powell and Tucker (2016) use the fact that the magnitude of the coefficient on *GDP Change for 1989* is larger when the corrected data is employed to suggest that the substantive effect of long-term economic performance is greater. This is incorrect. The magnitude of the coefficient on *GDP Change for 1989* is certainly larger with the corrected data, -6.00 [-21.1, 8.96]. The p -value associated with this coefficient is $p = 0.38$. The key point, though, is that the observed variation in *GDP Change for 1989* is now much smaller as the three miscalculated and outlier observations for Bosnia have been corrected. The mean value of *GDP Change for 1989* in the corrected data is 0.97 with a standard deviation of just 0.39; the minimum and maximum observed values are 0.22 and 2.04 (instead of 0.22 and 6.66). If we calculate the same two standard deviation swing as before, we now see a decrease in replacement volatility of just 4.69 [-15.48, 6.09]. Although the magnitude of the coefficient on *GDP Change for 1989* is 29.8% larger with the corrected data, one can see that the substantive effect of long-term economic performance on replacement volatility

is considerably smaller; in fact, it is just $\frac{4.69}{8.72} \times 100 = 53.8\%$ of what it was with the original incorrect data. We graphically compare the sizes of the two substantive effects in Figure 1. The substantive effect of a two standard deviation change in *GDP Change for 1989* on replacement volatility is much smaller with the corrected data and is estimated with much greater uncertainty.

Figure 1: The Substantive Effect of a Two Standard Deviation Increase in *GDP Change for 1989* on Replacement Volatility



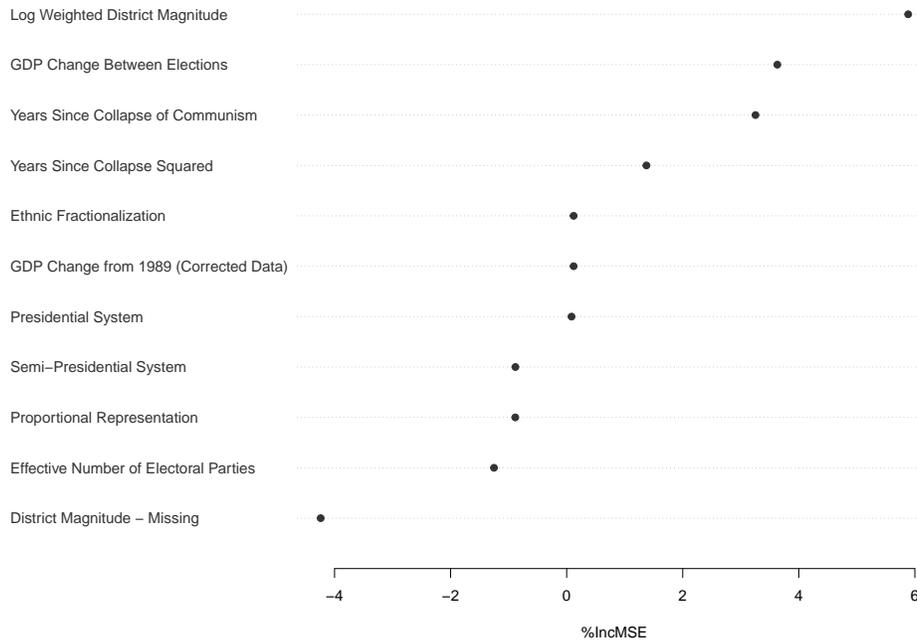
Note: Figure 1 shows the effect of a two standard deviation increase in *GDP Change for 1989* on replacement volatility using Powell and Tucker’s uncorrected data (top bar) and our corrected data (bottom bar). The horizontal bars represent two-tailed 95% confidence intervals.

While Powell and Tucker (2016) are wrong to suggest that the effect of long-term economic performance is substantively more important with the correct data, perhaps the absolute effect of long-term economic performance is nonetheless substantively important? This too, though, is incorrect. Replacement volatility has a mean of 31.86 and a standard deviation of 19.91; the minimum and maximum observed values are 1.45 and 82.09. A swing of two standard deviations in *GDP Change for 1989* using the corrected data leads to a change in replacement volatility that is smaller than 25% of replacement volatility’s standard

deviation. The substantive effect of *GDP Change for 1989* is small.

There are, of course, other ways to evaluate the substantive effect of a variable. None of these suggest that long-term economic performance has a substantively meaningful effect on replacement volatility in post-communist Europe either. In their original article, [Powell and Tucker \(2016, 134\)](#) write that they want to discover “predictors of . . . volatility.” As a result, we might want to evaluate whether *GDP Change for 1989* improves prediction and model fit. A simple test of this where we compare the R^2 and root mean squared error from a model that includes *GDP Change for 1989* ($R^2 = 0.109$, RMSE= 20.1) to the R^2 and RMSE from a model that excludes this variable ($R^2 = 0.102$, RMSE= 20.04) clearly indicates that this is not the case. Figure 2 shows a variance importance plot from a more sophisticated random forests model. Variables are ranked (top to bottom) in order of their importance for predicting replacement volatility. As Figure 2 indicates, there are five out of ten other variables (all statistically insignificant and

Figure 2: Variable Importance Plot for Powell and Tucker’s (2014) Model of Replacement Volatility



Note: Figure 2 plots the importance of the variables used in Powell and Tucker’s (2014) model of replacement volatility. On the vertical axis, variables are ranked (top to bottom) in order of their importance for predicting replacement volatility. The horizontal axis displays estimates of permutation accuracy for each variable. This is calculated as the difference in mean squared error between a model that is fitted using the real values for a measure and a model that is fitted using random (but realistic) values for the same measure. This measure is then scaled to represent the percentage increase in mean square error caused by permuting the values of the variable. Higher values indicate more important variables.

therefore rejected by Powell and Tucker as determinants of replacement volatility) that are considered more worthy of investigation in terms of their ability to predict replacement volatility than *GDP Change for 1989*.⁴ There is no compelling evidence, either statistically or substantively, to suggest that replacement volatility in post-communist Europe is driven by long-term economic performance.

Data Access and Research Transparency

Our research note has implications for the recent debate in political science about data access and research transparency (Golder and Golder, 2016). Data access and research transparency (DA-RT) has three components: (i) data access, (ii) production transparency, and (iii) analytic transparency. Data access indicates that researchers should reference the data they use to make evidence-based knowledge claims and provide access to it (when possible). Production transparency indicates that researchers should offer a full account of the procedures used to collect or generate the data. And analytic transparency requires that researchers show how they draw their analytic conclusions from the data.

In terms of data access, Powell and Tucker (2016, 2, note 6) claim that the “data on GDP in Bosnia-Herzegovina from 1989–93 . . . was unavailable at the time Powell and Tucker (2014) was submitted to the *British Journal of Political Science* for review,” which was “April 2011.” This is not true. While we cited the European Bank for Reconstruction and Development’s (EBRD) *Transition Report 2012* (Zettelmeyer, Ricka and Sanfey, 2012, 103) as the source for our data on Bosnia’s GDP in our research note, we were also able to find data for Bosnia’s 1989 GDP in the EBRD’s *Transition Report 2004* (Buiter and Fries, 2004, 38).⁵ This report and the data within it was available almost seven years prior to when Powell and Tucker submitted their initial manuscript for review.

We don’t want to make a big deal about this, though, as scholars often miss particular data sources. Indeed, as we point out in our research note, the choice to use one country’s 1994 GDP figure rather than its 1989 figure due to missing data “might ordinarily be expected to cause few problems” (Crabtree and Golder, 2016, 5). The more general issue here relates to production transparency. As anyone who has ever attempted to replicate someone else’s work knows, it is often difficult, if not impossible, to determine how particular variables are constructed or to identify specific data sources based solely on how these issues are

⁴Even if one were to use the uncorrected data, a variable importance plot from a random forests model reveals that four of the other ten variables are still considered more important for predicting replacement volatility than *GDP Change for 1989*.

⁵Although the EBRD’s *Transition Report 2004* does not provide Bosnia’s 1989 GDP figures directly, it can easily be calculated from the GDP figures on page 38 as 1989 is treated as the baseline year.

addressed in the published article or codebook. When it comes to DA-RT, it seems to us that the importance of analytic transparency has generally been taken as a given and that much of the debate has focused on the importance of data access. Relatively-speaking, production transparency has, unfortunately, been ignored.

To a large extent, issues with production transparency are not addressed by the recent practices that certain journals have introduced to check the replicability of research that has been accepted for publication. For example, the replication policy for the *American Journal of Political Science* is one of the most stringent in political science (<https://ajps.org/ajps-replication-policy/>). It requires that there be an in-house replication process before an article is published. The analysis conducted by Powell and Tucker (2014) would have passed this replication test as the replication data would have produced the results they report in their article. There would be no way to know that the values for *GDP Change from 1989* were not calculated as indicated for the three Bosnia observations, nor that a key substantive finding from their analysis depended on this fact. This suggests that attempts to improve data access and research transparency should encourage scholars not only to provide their ‘final’ dataset but also all of the original datasets and the code that was used to manipulate them into the ‘final’ dataset.

The only reason we became aware of this issue was because one of us, in the course of conducting a replication project for a quantitative methods course, used a bivariate plot to examine Powell and Tucker’s (2014) claim that replacement volatility was driven by long-term economic performance. This bivariate plot is shown in Figure 1 of our research note (Crabtree and Golder, 2016, 4). It was immediately obvious that the three Bosnia observations were distinct outliers. At this point, we checked to see if Powell and Tucker’s claim regarding the importance of long-term economic performance for replacement volatility was dependent on these outlier observations; it was. The extent to which the three Bosnia observations differed from the rest, though, made us wonder if the values were, in fact, correct. To see if this was the case, we tracked down the original data source (Pacek, Pop-Eleches and Tucker, 2009; Pop-Eleches, 2010) cited by Powell and Tucker (2014, 135) for the *GDP Change from 1989* variable and found that it did not contain any information on Bosnia’s GDP prior to 1994. Email correspondence with Powell and Tucker then revealed that they had not used 1989 as the reference year to calculate *GDP Change from 1989* for Bosnia.⁶ Instead, they had used 1994, the first available year for which they had data. Current practices with regard to replication could not have discovered any of this. Without producing the bivariate plot, we would have been

⁶As we mentioned in our research note (Crabtree and Golder, 2016, 1), but would like to reiterate here, we are grateful to Eleanor Powell, Joshua Tucker, and Grigore Pop-Eleches for the helpful comments they provided during the replication process.

none the wiser. As one of our reviewers explicitly noted, one of the most important contributions provided by our research note was to remind scholars of the value that comes from looking at simple descriptive statistics and plotting bivariate relationships before proceeding to more complicated models.

In their response, [Powell and Tucker \(2016, 2, note 8\)](#) seem to claim that in the end it doesn't really matter if they used 1989 or 1994, as they're not really sure what the most appropriate reference year for Bosnia is anyway. We think there are a number of problems with this. First, theoretical concepts need to be operationalized. If the decision has been made to use 1989 as the reference year when operationalizing one's theoretical concept, as was the case in [Powell and Tucker \(2014, 135\)](#), then it is inappropriate to not apply this operationalization to the three Bosnia observations as well. Second, if the theoretical concept being operationalized is the current level of economic performance in a country relative to what it was at the end of communist rule, as was the case in [Powell and Tucker \(2014, 135, 143\)](#), then it simply does not make sense to use 1994 as the reference year for Bosnia — communist rule was over by 1990 at the latest as this was when the “first post-communist multiparty elections in Bosnia and Herzegovina” were held ([Bieber, 2014, 548](#)). Third, if Powell and Tucker's theoretical concept involves comparing current economic performance to what it was when a country transitioned from communism, then using 1989 as a reference year for all of the observations in their sample makes little sense anyway. While 1989 represents the end of the communist period for some countries in Powell and Tucker's sample, like Bosnia-Herzegovina, several other countries were still communist in 1991. All of these issues suggest that not only is the effect of *GDP Change from 1989* on replacement volatility neither statistically nor substantively significant, but also that, given how it is operationalized, the *GDP Change from 1989* variable may not even be capturing a meaningful theoretical concept.

References

- ASA. 2016. "ASA Statement on Statistical Significance and *P*-Values." *The American Statistician* 70:131–133.
- Bieber, Florian. 2014. "Undermining Democratic Transition: The Case of the 1990 Founding Elections in Bosnia and Herzegovina." *Southeast European and Black Sea Studies* 14:548–555.
- Birch, Sarah. 2001. "Electoral System and Party System Stability in Post-Communist Europe." Paper presented at the 2001 Annual Conference of the American Political Science Association, San Francisco, CA, 26-29 August.
- Birch, Sarah. 2003. *Electoral Systems and Political Transformation in Post-Communist Europe*. Great Britain: Palgrave MacMillan.
- Buiter, Willem and Steven Fries, eds. 2004. *Transition Report 2004: Infrastructure*. London: European Bank for Reconstruction and Development.
- Crabtree, Charles and Matt Golder. 2016. "Party System Volatility in Post-Communist Europe." *British Journal of Political Science* .
- Golder, Matt and Sona N. Golder, eds. 2016. *Symposium: Data Access and Research Transparency (DART)*. CP: Newsletter of the Comparative Politics Organized Section of the American Political Science Association.
- Kitschelt, Herbert, Zdenka Manfeldova, Radoslaw Markowski and Gabor Toka. 1999. *Post-Communist Party Systems: Competition, Representation, and Inter-Party Cooperation*. Cambridge: Cambridge University Press.
- Pacek, Alexander C., Grigore Pop-Eleches and Joshua A. Tucker. 2009. "Disenchanted or Discerning: Voter Turnout in Post-Communist Countries." *Journal of Politics* 71:473–491.
- Pop-Eleches, Grigore. 2010. "Throwing Out the Bums: Protest Voting and Unorthodox Parties after Communism." *World Politics* 62:221–260.
- Powell, Eleanor Neff and Joshua A. Tucker. 2014. "Revisiting Electoral Volatility in Post-Communist Countries: New Data, New Results, and New Approaches." *British Journal of Political Science* 44:123–147.
- Powell, Eleanor Neff and Joshua A. Tucker. 2016. "Little is Known about Party System Volatility in Post-Communist Europe, but We Have Interesting New Methods and Data for Studying It." *British Journal of Political Science* .
- Tucker, Joshua A. 2006. *Regional Economic Voting: Russia, Poland, Hungary, Slovakia, and the Czech Republic, 1990-1999*. New York: Cambridge University Press.
- Wasserstein, Ronald L. and Nicole A. Lazar. 2016. "The ASA Statement on *P*-Values: Context, Process and Purpose." *The American Statistician* 70:129–131.
- Zettelmeyer, Jeromin, Franto Ricka and Peter Sanfey, eds. 2012. *Transition Report 2012: Integration Across Borders*. London: European Bank for Reconstruction and Development.