Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States

A Comment on Karadja and Prawitz (*Journal of Political Economy*, 2019)

Per Pettersson-Lidbom*

Journal of Comments and Replications in Economics, Volume 1, 2022-3, DOI: 10.18718/81781.25

JEL: D72; J61; P16
Keywords: Emigration; Omitted Variable Bias; Non-classical Measurement Error; Cluster-Robust Inference; Replication Study

Data Availability: The data and Stata code to reproduce the results of this comment can be downloaded at JCRE's data archive (DOI: 10.15456/j1.2022013.140613). The data and code of Karadja and Prawitz's paper discussed by Pettersson-Lidbom were published in the *Journal of Political Economy* (DOI: 10.1086/701682).

Please Cite As: Pettersson-Lidbom, Per (2022). Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States: A Comment on Karadja and Prawitz (2019). *Journal of Comments and Replications in Economics*, Vol 1(2022-3). DOI: 10.18718/81781.25

Abstract

In this comment, I revisit a question raised in Karadja and Prawitz (2019) concerning a causal relationship between mass emigration and long-run political outcomes. I find that their results are not robust to (i) selection of the appropriate control variables, (ii) using valid cluster-robust inference, (iii) using a more appropriate measure of emigration.

*Department of Economics, Stockholm University and Research Institute of Industrial Economics (IFN); Email: pp@ne.su.se

Acknowledgements: I am grateful to Erik Prawitz for providing a file with the names of the municipalities that enabled me to match my variable to the JPE data set. I am also grateful to Erik Prawitz, Mounir Karadja, Björn Tyrefors and David Strömberg for useful discussions.

Received January 13, 2022; Revised June 30, 2022; Accepted July 14, 2022; Published August 3, 2022.

1 Introduction

Karadja and Prawitz (2019) (KP) estimate the effect of emigration (i.e., external migration) on long-run political outcomes, such as labor movement membership, using historical data from Sweden for the period 1867–1920.

More specifically, KP estimate a cross-sectional regression in which the unit of observation is a municipality (i.e., 2359 geographical units). KP use an instrumental variable (IV) approach in which the instrumental variable for emigration is an interaction between a weather phenomenon, i.e., the number of frost shocks measured at the weather station level (i.e., 32 stations), and the geographical distance between the closest port of emigration and the municipality of residence. KP argue that their instrument is exogenous conditional on a set of control variables. They find strong support for the idea that emigration affects long-run political outcomes.

In this paper, I reinvestigate their claim. I have three problems with their analysis.[1]

The first concerns the selection of the appropriate set of control variables. I find that KP's results are extremely sensitive to which variables are included in the regressions. To circumvent difficulties with researcher degrees of freedom, I reanalyze KP's data using the post-double selection approach developed by Belloni et al. (2014). In sharp contrast with KP's result, I find no evidence that emigration has an effect on long-run political outcomes using the post-double selection method.

My second concern is whether the cluster-robust inference used by KP is reliable. Adopting the cluster-robust inference approach developed by Bell and McCaffrey (2002) and MacKinnon et al. (2022), I find that KP's results are driven by a small number of influential clusters with high leverage, which strongly indicates that KP's inference is not reliable. In fact, using a more reliable cluster-robust variance estimator suggests that KP's instrument is completely irrelevant.

My third concern is whether KP's measure of emigration is reliable. I find that KP's measure is unreliable because of a high proportion of missing data. When I use a more reliable measure, I again find that KP's instrument is irrelevant for explaining emigration.

The remainder of this paper is structured as follows. In Section 2, I discuss the question of how to select appropriate control variables. In Section 3, I analyze whether the cluster-robust inference used by KP is reliable. In Section 4, I discuss the problem of KP's measure of emigration not being reliable because of missing data. Section 5 concludes.

2 Selection of control variable

In this section, I investigate whether KP's analysis is sensitive to the selection of control variables. The reason for performing this specification test is that KP assume that their instrumental variable approach is only valid if the appropriate set of control variables is included in their specification.

[1] My critique also concerns Andersson, Karadja and Prawitz (2022), which uses the same identification strategy. My critique has previously been discussed in working papers, i.e., Pettersson-Lidbom (2020, 2022), which has been commented on by Kardaja and Prawitz (2020). However, Andersson, Karadja and Prawitz (2022) does not discuss my concerns.

This assumption is formally expressed on page 1886 of their paper as follows:

$$E[\varepsilon_{mct}|Shocks \times Port_{mc}, Shocks_{mc}, Port_{mc}, \Phi_c, \mathbf{X'}_{mc}] = 0. \tag{1}$$

Thus, KP's identifying assumption (i.e., the exclusion restriction) is that their instrument, i.e., *Shocks* × *Port*, only affects emigration conditional on a set of control variables, i.e., *Shocks*, *Port*, Φ_c (i.e., 24 county fixed effects), and other pretreatment variables as denoted by $\mathbf{X'}$.

As a starting point, I restate the baseline results from their paper, i.e., the first-stage results in Table 3 and the reduced-form effect on labor organization in Table 4. Column 1 in Panel A of Table 1 shows the results from the first stage of KP's analysis, while the reduced form appears in Column 1 in Panel B. The control variables included in these regressions are *Shocks*, *Port*, the *log of population in 1865*, and *county-fixed effects*. The first-stage estimate is 0.0635, with a standard error of 0.0157, and the estimated reduced form effect is 0.0014, with a standard error of 0.0004. Thus, both estimates are highly statistically significant, i.e., better than the 1% level. KP also show that these estimates are robust to the inclusion of additional covariates in their paper.

Surprisingly, however, KP do not present results from regressions without control variables in their tables, as is conventional in papers using quasi-experimental research designs. Instead, these results are displayed in Column 2 in Panels A and B in Table 1. Here, the first-stage estimate is 0.1396, with a standard error of 0.0723, while the estimated reduced form effect is −0.0004, with a standard error of 0.0010. Remarkably, none of the estimates are statistically significant at the 5% level, and the estimated reduced form effect is very different from the one reported by KP, i.e., −0.0004 versus 0.0014. To further probe the sensitivity of these results, I add the key control variables, *Shocks* and *Port*, in Column 3 since KP write, "An important feature of our identification strategy is that we control for the direct effects of frost shocks and port proximity". The first-stage estimate increases somewhat to 0.1646, with a standard error of 0.0597. However, the estimated reduced form effect remains the same, i.e., −0.0004. Thus, the inclusion of *Shocks* and *Port* makes little or no difference in the estimates. In Column 4, I also add the *log of population in 1865* to the regressions. Again, there are only small effects on the estimates. In sharp contrast, it is only when KP add region fixed effects (24 regions) to their specifications, as displayed in Column 1, that the first-stage and reduced form estimates change dramatically and become highly statistically significant.

There is however no compelling reason why one should control for geographical differences using Swedish counties, as I have argued in previous working papers Pettersson-Lidbom (2020, 2022). I have also discussed that a much more credible way to control for geographical characteristics is to include weather station fixed effects (32 stations) since such cluster-specific fixed effects enable one to completely control for any unmeasured (geographical) characteristics associated with KP's (clustered) instrument. In Column 5, I present the regression results when weather station fixed effects are included instead of county fixed effects. Strikingly, the reduced form effect becomes negative and statistically significant, i.e., −0.0015, with a standard error of 0.0006. This result is the opposite of that from KP's specification in Column 1. The estimated first stage is also no longer significant, i.e., 0.0421, with a standard error of 0.0316.

Table 1: Estimates of first-stage and reduced form effects

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. First-stage effects			
KP's instrument	0.0635	0.1396	0.1646	0.1890	0.0421	0.1040
	(0.0157)	(0.0723)	(0.0597)	(0.0413)	(0.0316)	(0.0360)
			Panel B. Reduced form effects			
KP's instrument	0.0014	−0.0004	−0.0004	−0.0002	−0.0015	−0.0006
	(0.0004)	(0.0010)	(0.0007)	(0.0006)	(0.0006)	(0.0011)

Notes: Column 1 includes the following control variables: *Shocks*, *Port*, the *log of population in 1865*, and *county-fixed effects*. Column 2 includes no control variables. Column 3 includes *Shocks* and *Port*. Column 4 includes *Shocks*, *Port*, the *log of population in 1865*. Column 5 includes*Shocks*, *Port*, the *log of population in 1865*, and weather station fixed effects. Column 6 includes control variables chosen by the post-double selection method as implemented by the Stata command `dsregress`. The dsregress is permitted to be selected from the universe of control variables (46) originally used by KP.

In summary, the results presented in Columns 1–5 strongly suggest that the first-stage and reduced form estimates are extremely sensitive to which control variables are included.[2] Unfortunately, this extreme sensitivity also opens up problems with respect to specification searches, researcher degrees of freedom and p-hacking (e.g., Leamer (1983), Simmons et al. (2011), and Brodeur et al. (2020)).

One way to address the problem of selecting the appropriate control variables is to use the post-double selection approach developed by Belloni et al. (2014). The idea behind post-double selection is to perform the selection in two steps: (i) find the control variables that predict the dependent variable and (ii) find the control variables that predict the independent variable. Belloni et al. show that the inclusion of both sets of control variables avoids overfitting and omitted variable biases.

Based on the post-double selection method, Column 6 in Table 1 presents results from the first-stage estimate (Panel A) and reduced form estimate (Panel B). Importantly, when performing post-double selection, I employ the same set of control variables (46) used by KP. While the post-double selection method shows that there is a statically significant first-stage estimate, i.e., 0.1040 (with a standard error of 0.0360), the estimated reduced form effect is close to zero and negative, i.e., −0.0006. These findings strongly suggest that the instrumental variable estimate should also be close to zero since there is no reduced form effect.

[2]In my working papers, Pettersson-Lidbom (2020, 2022), I present other specification tests that also reveal that KP's results are extremely fragile.

Table 2: Instrumental variable estimates

	(1)	(2)	(3)	(4)
KP's instrument	0.0213	0.0041	0.0015	0.0070
	(0.0075)	(0.0054)	(0.0017)	(0.0024)

Notes: Column 1 presents the result from KP's instrumental variable approach. Columns 2–4 present the results from different IV regressions that all use robust selection methods of control variables. These IV methods are all estimated with the Stata command ivlasso developed by Ahrens et al. (2020). Column 2 presents the result from the IV with PDS-selected variables and the full regressor set. Column 3 presents the results from the IV using the CHS postlasso-orthogonalized procedure. Column 4 presents the results from the IV using the CHS lasso-orthogonalized procedure.

To corroborate my claim, I have estimated instrumental variable specifications that also address the appropriate selection of control variables. Specifically, I use the `ivlasso` Stata command developed by Ahrens et al. (2020). Table 2 presents these results. In Column 1, I have restated the results from KP's IV specification with the most extensive set of controls, i.e., the results reported in Column 6 of Table 6. The IV estimate is 0.0231, with a standard error of 0.068. Column 2 presents the result from the IV with PDS (Post-Double-Selection) approach and the full regressor set. This IV estimate is 0.0041, with a standard error of 0.0054. Thus, the IV estimate is very small compared to that reported by KP, i.e., 0.0231. In Columns 3 and 4,I present the IV results from two other methods for the appropriate selection of control variables based on the "post-regularization" methodology of Chernozhukov, Hansen, and Spindler (2015): the CHS postlasso-orthogonalized procedure and the CHS lasso-orthogonalized procedure, respectively. Both these IV estimates are also very small, i.e., 0.0015 and 0.0070, respectively. Thus, and contrary to the claim in KP, there is no evidence of a causal effect of emigration on labor organization since all three IV estimates are close to zero.

3 *Clustered robust inference*

In this section, I analyze whether the cluster-robust inference used by KP is reliable. Specifically, I investigate whether the cluster-robust first-stage F tests reported in KP are valid. This analysis also provides a rationale for why the results of KP's analysis are extremely sensitive to which control variables are included, as discussed in the previous section.

The statistical inference problem arises because KP's data are clustered. For example, KP's instrument is constructed based on data from 32 weather stations. As a result, many observations contain the same weather data. For this reason, KP clusters their standard errors at the weather station level to address this issue. However, for the cluster-robust inference to be valid, the clusters must be both numerous and homogeneous, as discussed by Angrist and Pischke (2009) and Cameron and Miller (2015). Indeed, MacKinnon et al. (2022) show that cluster-robust inference is unreliable (i.e., severely over rejecting) if there are clusters that are very influential or with high leverage. In addition, they construct diagnostic tools to identify data sets and regression designs in

Table 3: Cluster variability: Leverage and partial leverage for the first-stage effect

Statistic	Cluster size	Leverage	Partial leverage
Minimum	2	0.18	0.0006
First quartile	11	0.71	0.0036
Median	65	1.57	0.0100
Mean	74	1.47	0.0312
Third quartile	107	1.97	0.0382
Maximum	311	4.77	0.1484
Coef. of variation	0.95	0.65	1.39

Notes: Notes: There are N=2359 observations and G=32 clusters. The effective number of clusters is $G^*_\gamma(0) = 11.1$ and $G^*_\gamma(1) = 5.9$.

which cluster-robust inference is likely to be challenging. Specifically, they compute a cluster-robust variance estimator that to perform much better in finite samples than other cluster-robust variance estimators. This cluster-robust variance estimator was originally developed by Bell and McCaffrey (2002).

I use the Stata command `summclust` developed by MacKinnon et al. (2022) to investigate whether the cluster-robust inference KP use is reliable. I start by presenting statistics on the degree of cluster heterogeneity, i.e., minimum, first quartile, median, mean, third quartile, and maximum of cluster sizes and leverage. Table 3 presents these results and shows that there is very large cluster heterogeneity. For example, the smallest cluster has only 2 observations, and the largest cluster has 311 observations. Moreover, both the leverage and the partial leverage vary considerably. The former ranges from 0.18 to 4.77, and the latter ranges from 0.0006 to 0.1484. The coefficients of variation are 0.65 and 1.39, respectively. The effective number of clusters (Carter et al. 2017) is also small, 5.9, compared to the original number of clusters, 32.

These results of very large cluster heterogeneity suggest that CV1, the default cluster robust variance estimator (CRVE), may not be particularly reliable in this case. MacKinnon et al. (2022) suggest that one should use an alternative CRVE, which is usually known as CV3, that has much better small sample properties. This estimator was originally proposed in Bell and McCaffrey (2002). Table 4 presents the regression results from CV1 and CV3. The results presented in the top row are the same as the first-stage estimate reported by KP (Column 6 in Table 3) since they also use CV1. If one instead uses CV3, the results are dramatically different since the first-stage estimate is no longer statistically significant (p-value=0.0745). This finding demonstrates that the cluster-robust inference used by KP is unreliable. One could also argue that KP do not use the correct clustering level since the instrument only takes 12 distinct values. Thus, the appropriate level on which to cluster should be 12, not 32. The final row in Table 4 presents CV3 at this level of clustering. Not surprisingly, this outcome indicates that the first-stage estimate is even more imprecisely estimated since the p-value is only 0.1487.

Table 4: Regression output for the first-stage estimate

s.e.	Coefficient	Standard Error	t-Statistic	p-value	CI lower bound	CI upper bound
CV1	0.062051	0.014660	4.2327	0.0002	0.032152	0.091950
CV3	0.062051	0.033622	1.8456	0.0745	−0.006521	0.130624
CV3	0.062051	0.039957	1.5530	0.1487	−0.025893	0.149995

Notes: The first row presents results from the default cluster robust variance estimator (CRVE) used by KP. The second and third rows present the results from the CRVE developed by Bell and McCaffrey (2002), which has much better small sample properties. The second row uses clusters based on weather stations (32 clusters), while the third row uses clusters based on the distinct values of the instrument (12 clusters).

The finding of very high degree of cluster heterogeneity also provides a rationale for why the results from KP's analysis are extremely sensitive to which control variables are included. In fact, the results presented previously suggests that one or two clusters drive all the results of KP, which leads to an extreme form of overfitting bias. This finding is related to the study by Young (2021), which shows that in this type of setting with highly clustered data the usual IV standard errors produced by Stata are susceptible to high leverage observations,[3] particularly with clustered and robust standard errors. In fact, he argues that "statistically significant IV results generally depend upon only one or two observations or clusters". Young proposes dropping one cluster at a time ("delete-one sensitivity") to check whether the statistical inference is reliable. Specifically, he argues that "delete-one sensitivity, of t-statistics not coefficients, highlights the degree to which significant results depend upon sensitive coefficient and standard error estimates". Applying the delete-one-sensitivity approach to KP's IV analysis reveals that dropping the cluster with three frost shocks produces a second-stage z-statistic of 1.37, with an associated p-value of 0.17.

4 *Measurement of emigration*

In this section, I discuss how KP's measure of emigration is unreliable due to missing data. I argue that it is better to use an alternative measure based on total outmigration, i.e., the sum of true emigration and internal migration since it (i) contains all emigrants (i.e., no missing data), (ii) has classical measurement error if KP's identifying assumptions are true (i.e., internal migration is not correlated with the instrument), and (iii) is robust to internal migration also being affected by KP's instrument.

It is well known that Swedish emigration statistics from the 19th century and early 20th century are unreliable due to the severe underreporting of emigrants. This fact has been documented and discussed, for example, by Statistics Sweden in an official report from 1887, Emigrationsutredningen (1913, p. 593), Johansson (1976), Odén (1964, 1971), Ahlqvist (1976), Eriksson et al. (1970), Hofsten and Lundström (1976), and Vernersson Wiberg (2016). These studies show not only that

[3]The standard errors produced by Stata are based on the fact that the normal approximation to the distribution of the reduced-form and first-stage coefficients is accurate. However, Young (2021) finds that normal approximation is unreasonable in settings with high leverage observations and clustered data.

the emigration to the U.S. was severely underreported but also that the emigration to other countries within Europe (e.g., Denmark and Germany) was even more underreported.

It is noteworthy that the studies that discuss the problems with Swedish emigration statistics are not cited in KP.[4] As a result, the argument in KP (p. 1876) that the emigration data are reliable is incorrect.[5] Specifically, the claim "it is possible to ascertain their [parish reports and ship passenger lists] accuracy by cross-checking the two sources" is erroneous since parish records reported emigration to all countries, while ship passenger lists essentially only recorded emigration to the United States. In fact, Eriksson (1970) finds that the overlap of individuals between these two sources is only 44%. Part of this discrepancy is due to parish records only registering individuals with change-of-address certificates.[6] Thus, KP cannot solve the underreporting problem by using a "single emigration variable defined as the maximum of either the church book or passenger list data each year" since there will be a very large number of missing emigrants.[7] Moreover, even unifying the two data sets would be insufficient since there would still be a large number of emigrants who are not recorded in either of these sources, i.e., those who did not apply for change-of-address certificates and who emigrated to countries other than the U.S. A similar point is made in Johansson (1976) and Odén (1971).

More importantly, I have estimated that KP's emigration variable only includes at most 73% of all emigrants during the period 1860–1920.[8] As a result of this large underreporting of emigration, the KP instrumental variable approach will be inconsistent due to a nonclassical measurement error, as discussed by Bound et al. (2001), who write (p. 3729) that "strategies for obtaining consistent estimates of the parameters of interest work if the measurement error is classical, but do not, in general do so otherwise".

To formally illustrate the measurement error problem in KP and how it can be resolved, I let X^* denote the true emigration. The population regression of interest in KP's analysis can now be expressed as follows:

$$Y_i = \alpha + \beta X_i^* + u_i, \tag{2}$$

where Y_i is some political outcome in municipality i, and X_i^* is the *true* total sum of emigrants who have emigrated (i.e., moved *outside* Sweden) from municipality i during the period 1867–1920.

[4]This literature should be familiar to KP since I provided references to it already in 2015 when I suggested that they must address the problem with measurement errors in the Swedish emigration statistics.

[5]KP's claim that their data sets encompass "the universe of registered emigrants during the Age of Mass Migration" is also incorrect since their parish data are estimated to contain only 75% of all emigrants. Data from a number of parishes are also missing from their data sets (see the following link: `emiweb.se/?services=emigranter-i-svenska-kyrkbocker/%20`).

[6]This problem has been regarded as the chief explanation of the discrepancy between actual and recorded emigration; Johansson (1976).

[7]Importantly, KP lack data from the church books after 1895.

[8]This calculation is partly based on official statistics (`www.scb.se/en/finding-statistics/statistics-by-subject-area/population/population-composition/population-statistics/pong/tables-and-graphs/yearly-statistics--the-whole-country/population-and-population-changes/`), i.e., the number of those with change-of-address certificates, which amounted to 1.3 million emigrants during the period 1860–1920. I have also estimated that a minimum of 0.2 million emigrants were not recorded during this period due to the various sources of error discussed by, e.g., Johansson (1976) and Eriksson et al.(1970). Thus, at least 1.5 million individuals emigrated from Sweden during the period 1860–1920. Consequently, a minimum of 0.4 million emigrants are missing from KP's data set since it only includes 1.1 million emigrants.

KP use an instrumental variable approach in which they replace the true value of X_i^* with an error-ridden measure, X_i, as noted above. Then, they assume that their instrumental variable, Z_i, is uncorrelated with both the population error term u_i and the reporting error $e_i = X_i - X_i^*$. However, because KP replace the true value in the equation with the error-ridden value, the instrument variable estimator is *not* consistent since the measurement error is not of the classical form due to underreporting.

The inconsistency problem in KP's instrumental variable approach can, however, be resolved by finding a measure of emigration that fulfills the assumption of classical measurement errors. In fact, the registered total outmigration, i.e., the sum of the true emigration, X_i^* and the true internal migration, I_i^*, fulfill the classical assumption since internal migration is *excluded* from the explanatory variables in KP's population regression model. That is, KP have (implicitly) assumed that their instrument Z_i is unrelated to internal migration I^* since it is subsumed in the population error term.[9] As a result, it is possible to replace X_i^* in equation (2) with total outmigration, i.e., $X_i^* + I_i^*$, and still obtain a consistent estimate of β since $Cov(I^*, Z)$ is assumed to be zero in KP's analysis.

An additional attractive feature of using total outmigration is that it is robust to the violation of KP's exclusion restriction, i.e., that internal migration is not affected by the instrument. To formally illustrate this point, note that KP's equation (1) states that

$$y_{cmt} = \beta Emigration_{cmt} + \Phi_S + X'_{mc}\beta_x + \eta_{cmt}. \tag{3}$$

However, a more general specification would also include internal migration as an explanatory variable:

$$y_{cmt} = \beta Emigration_{cmt} + \pi Internal\_migration + \Phi_S + X'_{mc}\beta_x + \eta_{cmt}. \tag{4}$$

In this case, two instruments would be required for identification since internal migration is an endogenous variable similar to emigration. Another solution to the identification problem is to redefine the parameter of interest as the effect of all types of migration, i.e., total outmigration, on the outcome. Thus,

$$y_{cmt} = \beta Total\_Migration + \Phi_S + X'_{mc}\beta_x + \eta_{cmt}. \tag{5}$$

where $Total\_Migration = Emigration + Internal\_migration$. In this case, it is sufficient to have only one instrument.

There are, in fact, good reasons to expect that internal migration is also affected by KP's instrument since KP show that internal migration is affected by the instrument (see Column 1 in Table 8 in KP). Thus, this result reveals that their instrument violates the exclusion restriction. Surprisingly, however, this issue is not discussed in the paper.

I have collected data on total outmigration from the Swedish National Archives for the period 1860–1950 as part of my ERC-financed historical database project. With these data, it is possible

[9] In fact, KP treat internal migration I^* as an additional outcome variable Y in their instrumental variable approach in Column 1 in Table 8. Thus, KP have assumed the following causal chain: $Z \to X^* \to Y$; i.e., the instrument Z only has an indirect effect, which only goes through X^*, on the outcome Y; i.e., $Cov(I^*, Z) = Cov(u, Z) = 0$. If this exclusion restriction is wrong, i.e., that Z has a *direct* effect on both X^* and I^*, then two valid instruments are required for identification, one for X^* and another for I^*. This may be another reason why KP's empirical analysis is flawed.

to assess the extent to which the results in KP are affected by the problem of underreporting of emigrants.[10] Interestingly, the reported emigration used by KP only makes up, on average, 8% of the total outmigration (the median value is 6%) during the period 1867–1920, and the share is never larger than 20% for any individual year. Thus, this value must be considered a very low share given the very large Swedish emigration during this period: it has been estimated that at least 1.5 million people emigrated, out of an average population of only 4.8 million, during the period 1860–1920. This finding further underscores the problem of underreporting Swedish emigration in KP's data.

If we turn to the result of the solution of the measurement error problem, the first-stage estimate with total outmigration is 0.010, with a standard error of 0.0076, while KP's estimate is 0.0621.[11] This first-stage estimate with total outmigration is also precisely estimated to be zero since it can rule out a first-stage effect larger than 0.026. Consequently, there is no first-stage relationship in KP's analysis when correcting for the problem of underreporting.

5 *Conclusion*

In this paper, I revisit the question raised in Karadja and Prawitz (2019) concerning a causal relationship between mass emigration and long-run political outcomes. I find that their results are not robust to (i) selection of the appropriate control variables, (ii) using valid cluster-robust inference, and (iii) using a more appropriate measure of emigration.

[10] In this footnote, I describe the result of merging my variable, the cumulative sum of total outmigration for the period 1867–1920, with KP's data set, which I downloaded from JPE's homepage. I discovered that KP's data files do not include the names of the geographical areas (i.e., municipalities) but only a variable running from 1 to 2,359. Thus, I had to ask KP to send me this information. After some work, I was able to match 2,330 out of the 2,359 municipalities by matching names.

[11] This specification corresponds to that appearing in Column 4 in Table 3 in KP, i.e., with a first-stage estimate of 0.621.

References

Ahlqvist, G. (1976). Sydsvensk utvandring till Danmark, Fallstudie av en i kyrkoböckerna registrerad utvandring, [Emigration from South Sweden to Denmark, A Case Study Based on the Registered Emigration from the Church Books]. Pdf: journals.lub.lu.se/scandia/article/view/1204/990.

Ahrens, A., C. B. Hansen, and M. E. Schaffer (2020). "lassopack: Model selection and prediction with regularized regression in Stata." *The Stata Journal*, 20(1):176-235. DOI: 10.1177/1536867X20909697.

Andersson, D., M. Karadja, and E. Prawitz (2022). "Mass Migration and Technological Change." *Journal of European Economic Association*, jvac017. DOI: 10.1093/jeea/jvac017.

Bell, R. M. and D. F. McCaffrey (2002). Bias reduction in standard errors for linear regression with multistage samples. *Survey Methodology*, 28: 169–81. Pdf: www150.statcan.gc.ca/n1/pub/12-001-x/2002002/article/9058-eng.pdf.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies*, 81(2): 608–50. DOI: 10.1093/restud/rdt044.

Brodeur, A., N. Cook, and A. Heyes (2020). "Methods Matter: p-Hacking and Publication Bias in Causal Analysis in Economics." *American Economic Review*, 110(11): 3634-60. DOI: 10.1257/aer.20190687.

Bound, J., C. Brown, and N. Mathiowetz (2001). "Measurement Error in Survey Data." In: J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Vol. 5, 3705-843. Amsterdam: Elsevier.

Cameron, A. C. and D. L. Miller (2015). "A practitioner's guide to cluster-robust inference." *The Journal of Human Resources*, 50(2): 317-72. DOI: 10.3368/jhr.50.2.317.

Carter, A. V., K. T. Schnepel, and D. G. Steigerwald (2017). "Asymptotic behavior of a *t*-test robust to cluster heterogeneity." *Review of Economics and Statistics*, 99(4): 698–709. DOI: 10.1162/REST_a_00639.

Chernozhukov, V., C. Hansen, and M. Spindler (2015). "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments." *American Economic Review*, 105(5): 486-90. DOI: 110.1257/aer.p20151022.

Emigrationsutredningen (1913). "Betänkande." Stockholm: Kungl. boktryckeriet. P. A. Norstedt & Söner.

Eriksson, I., K. Hvidt, C. Erickson, F. Nilsson, G. Rosander, L. Ljungmark, B. Brattne, and L. G. Tedebrand (1970). *Nordic Emigration: Research Conference in Uppsala Sept. 1969*. Uppsala: Soderstrom and Finn.

Hofsten E. and H. Lundström (1976). *Swedish Population History: Main Trends from 1750 to 1970 (Urval)*. Skriftserie Utgiven av Statistiska Centralbyran, Volume 8. Stockholm: National Central Bureau of Statistics.

Johansson R. (1976). "Registrering av flyttare, en källkritisk granskning av kyrkoboksmaterial 1840-90 [Registration of Migrants, a Source Critical Evaluation of the Church Books 1840-90]." *Scandia*, 42(2): 167-92. Pdf: journals.lub.lu.se/scandia/article/view/1203/989.

Karadja, M. and E. Prawitz (2019). "Exit, voice and political change: Evidence from Swedish mass migration to the United States." *Journal of Political Economy*, 127(4): 1864-925. DOI: 10.1086/701682.

Karadja, M. and E. Prawitz (2020). "A response to Pettersson-Lidbom's "Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States – a Comment." *Working Paper 2020:5*, Uppsala University, Department of Economics. Pdf: uu.diva-portal.org/smash/get/diva2:1457380/FULLTEXT01.pdf.

Leamer, E. E. (1983). "Let's Take the Con Out of Econometrics." *American Economic Review*, 73(1): 31–43. Pdf: www.jstor.org/stable/1803924.

MacKinnon, J. G., M. Ø. Nielsen, and M. D. Webb (2022). "Leverage, Influence, and the Jackknife in Clustered Regression Models: Reliable Inference Using summclust." *Working Paper 1483*, Economics Department, Queen's University. arXiv:2205.03288. DOI: arXiv.2205.03288.

Odén, B. (1964). "Den urbana emigrationen från Sverige 1840-1872 [The Urban Emigration from Sweden 1840-1872]." *Unpublished manuscript*, Lund.

Odén, B. (1971). "Ekonomiska emigrationsmodeler och historisk forskning. Ett diskussionsinlägg [Emigration Models in Economics and Historical Research: A Discussion]." *Scandia*, 37(1): 1-70. Pdf: journals.lub.lu.se/scandia/article/view/837/622.

Pettersson-Lidbom, P. (2020). "Exit, Voice and Political Change: Evidence from Swedish Mass migration to the United States: A Comment." *Research Papers in Economics 2020:3*, Stockholm University, Department of Economics. Pdf: www2.ne.su.se/paper/wp20_03.pdf.

Pettersson-Lidbom, P. (2022). "Exit, Voice and Political Change: Evidence from Swedish Mass migration to the United States: A Comment." *Working Paper, Version: January 12, 2022*. DOI: arXiv.2201.04880.

Simmons, J. P., L. D. Nelson, and U. Simonsohn (2011). "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science*, 22(11): 1359–66. DOI: 10.1177/0956797611417632.

Vernersson Wiberg, A.-K. (2016). "Migration och Identitet: En studie av utvandringen från Blekinge till Danmark och Tyskland 1860–1914. [Migration and Identity: A Study of Emigration from Blekinge to Denmark and Germany 1860-1914]". *Licentiate Thesis*. Uppsala: Department of Social and Economic Geography, Uppsala University. Pdf: www.diva-portal.org/smash/get/diva2:925667/FULLTEXT02.pdf.

Young, A. (2021). "Leverage, Heteroskedasticity and Instrumental Variables in Practical Application." *Unpublished Working Paper*. London School of Economics.