

Akhtar, Farida; Kreitmeir, David; Newman, Luke; Ploeckl, Florian; Tarlinton, Boyd

**Working Paper**

## A Comment on "Winter is Coming: Early-Life Experiences and Politicians' Decisions"

I4R Discussion Paper Series, No. 155

**Provided in Cooperation with:**

The Institute for Replication (I4R)

*Suggested Citation:* Akhtar, Farida; Kreitmeir, David; Newman, Luke; Ploeckl, Florian; Tarlinton, Boyd (2024) : A Comment on "Winter is Coming: Early-Life Experiences and Politicians' Decisions", I4R Discussion Paper Series, No. 155, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/302899>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



No. 155

I4R DISCUSSION PAPER SERIES

# **A Comment on “Winter is Coming: Early-Life Experiences and Politicians’ Decisions”**

Farida Akhtar

Luke Newman

Boyd Tarlinton

David Kreitmeir

Florian Ploeckl

September 2024

## I4R DISCUSSION PAPER SERIES

I4R DP No. 155

### **A Comment on “Winter is Coming: Early-Life Experiences and Politicians’ Decisions”**

**Farida Akhtar<sup>1</sup>, David Kreitmeir<sup>2</sup>, Luke Newman<sup>2</sup>, Florian Ploeckl<sup>3</sup>,  
Boyd Tarlinton<sup>4</sup>**

*<sup>1</sup>Macquarie University, Macquarie Park/Australia*

*<sup>2</sup>Monash University, Melbourne/Australia*

*<sup>3</sup>University of Adelaide/Australia*

*<sup>4</sup>Queensland Department of Agriculture and Fisheries, Brisbane/Australia*

SEPTEMBER 2024 (update of the former September 2024 version)

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

#### **Editors**

**Abel Brodeur**  
*University of Ottawa*

**Anna Dreber**  
*Stockholm School of Economics*

**Jörg Ankel-Peters**  
*RWI – Leibniz Institute for Economic Research*

# A comment on “Winter is Coming: Early-Life Experiences and Politicians’ Decisions”\*

Farida Akhtar      David Kreitmeir      Luke Newman  
 Florian Ploeckl      Boyd Tarlinton

August 08th, 2024

## Abstract

Guo et al. (2023) examine the impact of early-life experiences of politicians on their policy implementations. They utilize differential exposure of Chinese county party secretaries to the Great Famine of 1959-1961 as a natural experiment and investigate the impact on their policy preferences, in particular fiscal expenditure on agriculture and social security. In their baseline analytical specification, the authors find that exposure to a one percentage point more severe famine led counties governed by these politicians to a 0.8% higher fiscal expenditure on agriculture and a 1.1% higher expenditure on social security subsidies. Their point estimates are statistically significant at the 1%, respective 5% levels depending on the included set of controls.

First, we successfully computationally reproduce all quantitative claims, more precisely all tables and figures, of the paper, using the provided replication files. We uncover a minor coding error in a specification in a robustness check, though correcting it does actually strengthen the studies’ main result, as well as a typo and rounding error in another robustness check. Additionally, the summary statistics and exploratory data analysis of the paper were also computationally reproduced using a different software package.

Second, we directly replicate the results by systematically varying the sample size in two ways. One, we drop one individual control variable which increases the sample by retaining the observations that have missing values for that control variable, and two, we restrict the sample for the estimation of the impact on social security subsidies to only those observations that also report values for the agricultural fiscal subsidies.

We find that retaining the observations with missing educational information values reduces the magnitude of all coefficients of interest (so interactions of famine severity with birth year indicators) with the impact on agricultural expenditure remaining statistically significant while the impact on social security subsidies is no longer statistically significant in three specifications and

---

\*Authors: Akhtar: Macquarie University. E-mail: [farida.akhtar@mq.edu.au](mailto:farida.akhtar@mq.edu.au). Kreitmeir: Monash University. E-mail: [david.kreitmeir1@monash.edu](mailto:david.kreitmeir1@monash.edu). Newman: Monash University. E-mail: [luke.newman1@monash.edu](mailto:luke.newman1@monash.edu). Ploeckl: University of Adelaide. E-mail: [florian.ploeckl@adelaide.edu.au](mailto:florian.ploeckl@adelaide.edu.au). Tarlinton: Queensland Department of Agriculture and Fisheries. E-mail: [boyd.tarlinton@daf.qld.gov.au](mailto:boyd.tarlinton@daf.qld.gov.au)

For each author: No conflict of interest to report. No financial support to acknowledge. See I4R’s conflict of interest policy here: <https://i4replication.org/conflict.html>.

remains weakly significant in one. Similarly, the impact of estimating the impact on social security subsidies with the restricted sample of observations that have agricultural subsidy values also reduces the magnitude substantially and turns the coefficient statistically insignificant.

## 1 Introduction

[Guo et al. \(2023\)](#), henceforth GGL, report on links between famine experienced during the childhood of Chinese county party secretaries and their later political activities. They investigate the fiscal expenditure of Chinese counties on agriculture and social security in the period 1998 to 2007 and find that counties governed by party secretaries with a stronger exposure to early-life famine environments have higher expenditures in both areas. They ascribe the higher expenditures to the impact on preference formation by those early-life experiences of material deprivation.

GGL test the impact of the famine experience using a difference-in-difference approach comparing the expenditure in counties that see a party secretary with early-life famine experience come to power to counties governed by party secretaries without that experience. The data set is a combination of published statistical information about Chinese counties with information collected about individual party secretaries' life course. Only counties with sufficient background information about their party secretaries are included in the sample.

GGL describe their main result on page 297 as follows: “We find that, for County Party Secretaries (CPSs) who experienced famine in early life (i.e., born 1949–60), if the famine that affected their birth counties was more severe by one percentage point, fiscal expenditures on agriculture in the counties they governed increased by 0.8% and expenditures on social security subsidies increased by 1.1% when compared with expenditures in counties governed by CPSs who were born after the famine. The mean of our famine-severity variable is 0.43 (i.e., a drop of 43 percentage points in county-level agricultural populations born during famine years). Hence, an increase in famine intensity by 1% of the mean corresponds to an increase in agricultural spending of about 0.4% and an increase in social security subsidy spending of 0.5%.”

In the present report prepared for the Institute for Replication ([Brodeur et al. 2024](#)), we investigate whether their analytical results are computationally reproducible and further test their replicability and robustness to changes in the sample composition used for the analysis. We test two alternative sample compositions using the provided paper data set: First all specifications are estimated without the “eduyear” control variable. As a number of observations were excluded by GGL

from the estimation due to missing values about the education of the party secretaries, as captured in the "eduyear" variable, this increases the sample size for all main specifications. Second, we estimate the specifications using social security subsidies only for observations with agricultural expenditure values available. This leads to a substantial reduction in included observations due to the high number of missing agricultural expenditure values.

In terms of reproducibility, first we would like to acknowledge that the team at the Economic Journal successfully reproduced the results of the original study with the code provided by the authors and submitted to the relevant depository. We also successfully reproduced GGL's results throughout the whole paper. However, we uncovered one minor coding error involving the difference between birth and position prefectures (Table C5) as well as a small number of additional discrepancies, which we ascribe to typos or rounding differences.

We were able to reproduce exploratory figures and summary statistics in R without reference to Stata code. Additionally, we were able to reproduce the main results as depicted in Tables 2 through 4 with reference to the Stata code.

In terms of robustness and sensitivity analysis, we estimate the main specifications with a different selection of the included observations.

Removing the "eduyear" variable does not have any noticeable impact on the result of the original specification when estimated with the same sample. Dropping the variable from the main specifications and including the observations with missing "eduyear" values increases the sample size by 1011 (1146) for agricultural expenditure specifications and 1698 (1930) for social security subsidy ones. The resulting coefficients for the main variable of interest, the interaction between birthcohort and famine intensity, reduce substantially in magnitude (from 0.89 to 0.58 for the preferred agricultural specification and from 1.06 to 0.18 for the social security one). The coefficients remain statistically significant for the agricultural expenditure specifications but are no longer significant in three social security subsidy specifications and weakly significant at the 10% level in one.

The second analysis harmonizes the samples between the specifications involving agricultural expenditure as outcome variables and those covering social security subsidies. Practically this means that we estimate the specifications for social secu-

rity subsidies only with those observations with available agricultural expenditure values. As a result the coefficients turn insignificant and substantially smaller for specifications involving the full affected cohort, while the coefficients for the 49-54 cohort turn substantially and statistically significantly negative though the effects for the 55-60 cohort only turns insignificant for one specification.

## 2 Computational Reproducibility

We used the replication package provided: <https://doi.org/10.5281/zenodo.8188982>. The package does not contain raw data and cleaning codes, but only the analysis data and respective code.

Table 1: Replication Package Contents and Reproducibility

Replication Package Item	Fully	Partial	No
Raw data provided			✓
Analysis data provided	✓		
Cleaning code provided			✓
Analysis code provided	✓		
Reproducible from raw data			✓
Reproducible from analysis data	✓		

*Notes:* This table summarizes the replication package contents contained in the replication package linked by Guo et al. (2023).

We successfully computationally reproduced all results presented in the paper using the provided data and code files. Instructions were sufficiently detailed and all required software packages were available and downloadable from the indicated locations.

To keep consistency in reporting, we uncovered some very minor discrepancies which we consider to be typing or rounding errors:

For Table B4 Panel A Column 6 our results show 2.28 (1.35) instead of 2.29 (1.28) in the paper.

For Table C1 Panel B Column 5 + 6 our results show for the R-square 0.13/0.19 instead of 0.14/0.20 in the paper.



For Table C7 Column 1 our results show for the R-square a value of 0.0311 instead of 0.31 in the paper

Additionally, we uncover an inconsequential coding error in a robustness specification. Columns 7 and 8 in Table C5 in the online appendix include an interaction of the main variable of interest with a dummy variable indicating that the home county and the current county of a party secretary are located in different prefectures. In contrast to the description, the coding in the replication file coded the variable as 1 when the two counties were in the same prefecture. Modifying the variable and re-estimating the specifications results in coefficients that are much more in line with the results of the other robustness specifications included in Table C5. The re-estimated coefficients are reported in Table 2

### 3 Recreate Reproducibility

To assess the completeness of the methods described in the paper, we attempted to reproduce the analyses in R 4.3.3 without first viewing the Stata code. We recreated all summary statistics and figures from the described methods without incident. We were further able to recreate the key results contained in Tables 2 through 4 with some reference to the Stata code in order to determine the correct specifications for some models. The results of our recreated scripts conform with those presented in the original paper, noting that the incorrectly specified variable noted in the computational reproducibility section was again present in the data.

### 4 Robustness Replication: Sample Selection

We now turn our attention to our main replication. We test the robustness of the results to direct replications with differently selected sub-samples of the full data set.

The decision to conduct these two robustness checks was taken only after conducting the computational replication described above. consequently, the analysis has not been pre-registered. During the computational replication, we noticed the

notable differences in included observations between the summary statistics (8,873 in GGL Table 1), the agricultural specifications in Table 2 (4,020/3,532), and the social security specifications in Table 2 (6,655/5,872). We explored the origins of these differences and the consequences of the different sample sizes and found the two important aspects presented below.

For comparison purposes Table 6 is a direct replication of Table 2 in GGL with p-values rather than standard errors.

#### 4.1 Educational Attainment control variable with missing values

As part of the data sample exploration we were exploring why the estimated specifications use substantially less observations than available in the overall data set. We identified the "eduyear" control variable, which contains the years of formal education a party secretary had. The variable was collected by GGL from biographies of the politicians and shows a substantial number of missing values.

The impact of these missing values is such that for the main agricultural specifications ( GGL Table 2) 1011 (specifications with county covariates) respective 1146 (specifications without county covariates) observations are dropped solely due to missing educational attainment values. For social security subsidy specifications 1698 observations are dropped for specifications with county covariates and 1930 observations for specifications without such covariates.

We estimated agricultural and social security specifications (columns 1 and 5 of Table 2) without the "eduyear" variable. Table 3 compares the results. There is effectively no impact on the agricultural specification, while educational attainment is statistically significant for the social security subsidy specification. However, the magnitude of the effect is small compared to the impact of the difference in sample size, so it cannot fully explain the results presented below.

Table 4 presents the results of re-estimating the specifications included in Table 2 of GGL with the difference that the "eduyear" control variable has been removed and the observations with missing educational attainment variables have been added.

For specifications involving agricultural expenditures, the coefficients retain statistical significance, however, their magnitude drops by about 30% over all four specifications. For specifications involving social security subsidies, the coefficients

of three specifications become statistically insignificant and one retains weak significance at the 10%-level. The magnitude drops by 54% – 83% and gets substantially closer to zero.

## 4.2 Impact of missing agricultural expenditure values

The second notable observation about the number of observations in GGL Table 2 is the difference between the number included in the agricultural expenditure specifications (4,020/3,532) and those included in the corresponding social security subsidy specifications (6,655/5,872).

To explore the sensitivity of the results to that difference we re-estimate the social security subsidy specifications with the samples used for the agricultural expenditure specifications. There is also a very small number of observations that have missing social security subsidy values but are not missing agricultural expenditure values, so we end up using 3998 instead of 4020 and 3524 instead of 3532 observations.

The results corresponding to columns 5 and 6 of GGL Table 2 show that the coefficient magnitude again dropped substantially and both coefficients are statistically insignificant. The results for columns 7 and 8 even show a substantially negative effect for the 1949-1954 cohort, which is statistically significant at 5% respective 10%-level, and notably negative effect. The effect for the 1955-1960 cohort remains positive and weakly significant at a 10%-level without county controls but turns statistically insignificant with county controls.

## 5 Conclusion

This analysis undertakes a replication of [Guo et al. \(2023\)](#). We find the code to well documented and were able to conduct a computational replication without significant problems or differences. A minor coding error in a robustness check even strengthens the robustness of the results.

A notable issue, however, is the sensitivity and robustness of the main results to the utilized samples of observations. This is especially a problem for the results of the specifications explaining the impact of early-life famine experience on social

security subsidies.

The inclusion of observations with missing educational attainment reduces the magnitude of the impact on agricultural expenditure, but retains the statistical significance. The impact on social security subsidies, however, becomes statistically insignificant with a substantially diminished magnitude. Similarly, the estimation of the impact on social security subsidies with observations that also have agricultural expenditure values again severely reduces coefficient magnitude and leads to statistical insignificance.

First explorations why these particular changes to the utilized sample result in such different outcomes have not yet been conclusive.

### **Replication code**

The code used in this replication is publicly available at <https://zenodo.org/records/13710456>.

## References

- Brodeur *et al.*: 2024, Mass Reproducibility and Replicability: A New Hope, *I4R Discussion Paper Series* .
- Guo, S., Gao, N. and Liang, P.: 2023, Winter is Coming: Early-life Experiences and Politicians' Decisions, *The Economic Journal* **134**(657), 295–321.  
**URL:** <https://doi.org/10.1093/ej/uead061>

## 6 Tables

Table 2: Table C5 correction

VARIABLES	(7) Agri	(8) Soc Sec
c.famineseverity#c.born4960	0.70** (0.30)	0.82* (0.46)
c.famineseverity#c.born4960#c.diffcity	0.78 (6.37)	1.49 (1.23)
Observations	3,532	5,872
Within R-squared	0.0285	0.0262

Robust standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Notes: Estimated Coefficients when "diffcity" is coded as described in the online appendix of [Guo et al. \(2023\)](#).

Table 3: Comparison Dropping "eduyear" control variable

VARIABLES	(1) Agriculture	(2) Agriculture	(3) SocialSecurity	(4) SocialSecurity
c.famineseverity#c.born4960	0.89*** (0.30)	0.89*** (0.31)	1.01** (0.43)	1.07** (0.44)
eduyear	0.03 (0.03)		0.07** (0.03)	
c.eduyear#c.born4960	-0.00 (0.04)		-0.07* (0.04)	
Constant	16.49*** (1.01)	17.01*** (0.96)	15.63*** (1.13)	16.08*** (1.10)
Observations	4,020	4,020	6,655	6,655
R-squared	0.95	0.95	0.91	0.91

Robust standard errors in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Notes: Columns 2 and 4 are identical to columns 1 and 3 except for the dropped "eduyear" control variable..

Table 4: Main Specifications - "Eduyear" variable dropped

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Agricultural expenditure			Social security subsidies			
c.famineseverity#c.born4960	0.58*** (0.01)	0.58** (0.01)			0.37 (0.24)	0.18 (0.56)		
c.famineseverity#c.born4954			0.33 (0.25)	0.30 (0.40)			-0.27 (0.50)	-0.30 (0.42)
c.famineseverity#c.born5560			0.66*** (0.00)	0.66*** (0.01)			0.68* (0.05)	0.42 (0.21)
County Covariates		Yes		Yes		Yes		Yes
CPS Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,166	4,543	5,166	4,543	8,585	7,570	8,585	7,570
R-squared	0.94	0.94	0.94	0.94	0.91	0.91	0.91	0.91
Number of work counties	1336	1170	1336	1170	1554	1349	1554	1349
Adjusted R-squared	0.897	0.881	0.897	0.881	0.865	0.869	0.866	0.869
Within R-squared	0.00230	0.0241	0.00247	0.0243	0.00138	0.0271	0.00255	0.0277

Robust p-values in parentheses, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Notes: Authors' calculations using GGL data set with each specification modified to drop "eduyear" control variable. Observations with missing "eduyear" values are included. Table reports p-values.

Table 5: Sample with existing agricultural expenditure values

VARIABLES	Social Security Subsidies			
	(5)	(6)	(7)	(8)
c.famineseverity#c.born4960	0.52 (0.43)	0.28 (0.70)		
c.famineseverity#c.born4954			-2.25** (0.03)	-1.87* (0.06)
c.famineseverity#c.born5560			1.17* (0.09)	0.79 (0.28)
Observations	3,998	3,524	3,998	3,524
R-squared	0.92	0.92	0.92	0.92
Number of work counties	1104	972	1104	972
Adjusted R-squared	0.838	0.838	0.838	0.838
Within R-squared	0.00207	0.0214	0.00453	0.0228

Robust p-values in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Notes: Authors' calculations using GGL specifications and data set modified to only include observations with agricultural expenditure values in the estimation. Table reports p-values



Table 6: computational Replication of GGL Table 2 - P-values displayed

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Agricultural expenditure				Social security subsidies		
c.famineseverity#c.born4960	0.89*** (0.00)	0.83*** (0.00)			1.01** (0.02)	1.06*** (0.01)		
c.famineseverity#c.born4954			0.66 (0.17)	0.35 (0.45)			-0.05 (0.93)	0.43 (0.38)
c.famineseverity#c.born5560			0.94*** (0.00)	0.95*** (0.00)			1.48*** (0.00)	1.35*** (0.00)
County Covariates		Yes		Yes		Yes		Yes
CPS Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,020	3,532	4,020	3,532	6,655	5,872	6,655	5,872
R-squared	0.95	0.93	0.95	0.93	0.91	0.92	0.91	0.92
Number of work counties	1107	972	1107	972	1335	1164	1335	1164
Adjusted R-squared	0.893	0.867	0.893	0.867	0.865	0.868	0.865	0.868
Within R-squared	0.00352	0.0273	0.00359	0.0276	0.00426	0.0278	0.00612	0.0284

Robust p-values in parentheses

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Notes: Authors' calculations using GGL data set and GGL code. Table reports p-values instead of standard errors.