![Gmail]

Frederik Jølving <fcjoelving@gmail.com>

journalist inquiry
11 messages

Frederik Joelving <fcjoelving@gmail.com>   Fri, Feb 2, 2024 at 10:56 AM
To: almas.heshmati@ju.se, almas.heshmati@gmail.com

Dear Professor Heshmati,

I'm a journalist with Retraction Watch, a New York-based media organization, and would like to talk to you about your
paper "Green innovations and patents in OECD countries" for a story I'm working on . If you could give me a call
today, I'd appreciate it very much.

Thanks,
Fred

--
Frederik Joelving
Journalist
Editor, Retraction Watch

+45 7172 1423
http://www.joelving.com/

Almas Heshmati <Almas.Heshmati@ju.se>   Fri, Feb 2, 2024 at 11:09 AM
To: Frederik Joelving <fcjoelving@gmail.com>

Dear Frederik,

Am sorry for not answering your phone call. I did not recognize the number. I am changing my phone with the help of
IT helpdesk. Have no phone now. Is it possible to ask your questions by e-mail. I promise to answer then as quick as
possible. I write many papers and need to go back to the paper to answer your questions.

Best regards,

Almas

Almas Heshmati

Professor of Economics,

Jönköping International Business School (JIBS),

Jönköping University, Room B5017,

Gjuterigatan 5, SE-551 11 Jönköping, Sweden,

Phone: +46-36-101780, Cell: +46-739-101877,

E-mail: almas.heshmati@ju.se

[Quoted text hidden]

Frederik Joelving <fcjoelving@gmail.com>   Fri, Feb 2, 2024 at 11:39 AM
To: Almas Heshmati <Almas.Heshmati@ju.se>

Thank you, Almas. Please find my questions below:

1) It has come to my attention that the data you relied on for your paper has a lot of gaps. But you say you used balanced panel data in the paper. Can you please explain?

2) You provide descriptive statistics based on "783 observations" for each country in Table 1. But there are a lot of data gaps for each variable, so isn't it misleading to say there were 783 observations?

3) It appears you based the descriptive statistics on imputed data. Shouldn't that have been mentioned somewhere in the article?

4) How did you do the imputations and in which program?

5) Did you do other operations than imputations to fill in the missing data?

6) Have you done any sensitivity analyses to see what happens if you don't fill out the data gaps?

7) How many data points do you think you imputed and why do you not mention this in your paper?

8) Was your late coauthor, Mike Tsionas, aware of the way you dealt with missing data?

Thanks,
Fred
[Quoted text hidden]

Frederik Joelving <fcjoelving@gmail.com>   Fri, Feb 2, 2024 at 12:42 PM
To: Almas Heshmati <Almas.Heshmati@ju.se>

Sorry, Question 2 should have read:

2) You provide descriptive statistics based on "783 observations" for each variable in Table 1. But there are a lot of data gaps for each variable, so isn't it misleading to say there were 783 observations?
[Quoted text hidden]

Almas Heshmati <Almas.Heshmati@ju.se>   Fri, Feb 2, 2024 at 1:28 PM
To: Frederik Joelving <fcjoelving@gmail.com>

Dear Fredrick,

Your questions are not friendly and accusing. Please see my answers below the questions.

Best regards,

Almas

Almas Heshmati

Professor of Economics,

Jönköping International Business School (JIBS),

Jönköping University, Room B5017,

Gjuterigatan 5, SE-551 11 Jönköping, Sweden,

Phone: +46-36-101780, Cell: +46-739-101877,

E-mail: almas.heshmati@ju.se

From: Frederik Joelving <fcjoelving@gmail.com>
Sent: Friday, 2 February 2024 12:43
To: Almas Heshmati <Almas.Heshmati@ju.se>
Subject: Re: journalist inquiry

Sorry, Question 2 should have read:

2) You provide descriptive statistics based on "783 observations" for
each variable in Table 1. But there are a lot of data gaps for each
variable, so isn't it misleading to say there were 783 observations?

Answer: The final data used has no missing or gaps. Missing observations are imputed using forward and backward
trend imputation based on the same countries' variables. This will not change the result but fill the gaps which is
necessary when one use dynamic model with lags and changes.

On Fri, Feb 2, 2024 at 11:39 AM Frederik Joelving <fcjoelving@gmail.com> wrote:

> Thank you, Almas. Please find my questions below:
>
> 1) It has come to my attention that the data you relied on for your
> paper has a lot of gaps. But you say you used balanced panel data in
> the paper. Can you please explain?
>
> Answer: The final data used has no missing or gaps. Missing observations are imputed using forward and
> backward trend imputation based on the same countries' variables. This will not change the result but fill the gaps
> which is necessary when one use dynamic model with lags and changes.
>
> 2) You provide descriptive statistics based on "783 observations" for
> each country in Table 1. But there are a lot of data gaps for each
> variable, so isn't it misleading to say there were 783 observations?
>
> Answer: The final data used has no missing or gaps. Missing observations are imputed using forward and
> backward trend imputation based on the same countries' variables. This will not change the result but fill the gaps
> which is necessary when one use dynamic model with lags and changes.
>
> 3) It appears you based the descriptive statistics on imputed data.
> Shouldn't that have been mentioned somewhere in the article?

Answer: Imputations was mainly for Island in early years.

4) How did you do the imputations and in which program?

Answer: using forward and backward trend based on 3 observation for same country and variable.

5) Did you do other operations than imputations to fill in the missing data?

Answer no: You can use imputation in STAT but do not recommend it. My approach since it is trend based will not change the result. Unit imputation is necessary not to loss the whole observation and all other variables for a given country and year.

6) Have you done any sensitivity analyses to see what happens if you don't fill out the data gaps?

Answer: no but I think trend does not change the result unless in cases of countries and period with radical changes.

7) How many data points do you think you imputed and why do you not mention this in your paper?

Answer: I do not recall, if not written there must have forgotten.

8) Was your late coauthor, Mike Tsionas, aware of the way you dealt with missing data?

Answer: yes

[Quoted text hidden]

[Quoted text hidden]

Frederik Joelving <fcjoelving@gmail.com>   Fri, Feb 2, 2024 at 1:55 PM
To: Almas Heshmati <Almas.Heshmati@ju.se>

Dear Almas,

I appreciate your taking the time to answer my questions. Please allow me a few follow-up questions:

1) You do not address why you do not mention the imputations in the paper. Should this not have been done?

2) Does "observation" not refer to a measured variable rather than a calculated one in economics?

3) I have had a chance to look at your data in Excel. There are not just a few, but **thousands** of imputed values that do not appear in the raw data. In addition, you say you did not perform other operations than imputation to fill the missing data gaps. However, I can see that in 5 cases, you have copied data from one country to another country, and these copied data are included in your analysis. Experts I have consulted about this have expressed alarm at such operations. They have also said that not describing these operations in the paper amounts to "cheating." Do you have any comments on this?

Sincerely,
Fred
[Quoted text hidden]

Almas Heshmati <Almas.Heshmati@ju.se>   Fri, Feb 2, 2024 at 2:02 PM
To: Frederik Joelving <fcjoelving@gmail.com>

Dear Fredrik,

You are harsh in your tone and interrogating. I have no obligation to answer your e-mail yet try to be kind and communicate. How about introducing yourself first and what you do and why you question my research. If we do not use imputation, such data is almost useless in content and coverage as in many rows there is a unit of missing eliminating the observation from estimation.

The mention of imputation left out not been intentional. Tell me how you got hold of the original data and my file used in the estimation? We have not been in touch before.

[Quoted text hidden]

Frederik Joelving <fcjoelving@gmail.com>   Fri, Feb 2, 2024 at 2:11 PM
To: Almas Heshmati <Almas.Heshmati@ju.se>

Dear Almas,

I did introduce myself, and my email signature contains links to my organization's website as well as to my personal site. I cannot tell you how I got the data, but the fact that your paper does not mention anything about imputations or copied data does seem to compound the problems inherent in those operations. I you have additional comments, please let me know. I plan to file my story early next week.

Thanks,
Fred

.
[Quoted text hidden]

Almas Heshmati <Almas.Heshmati@ju.se>   Fri, Feb 2, 2024 at 2:15 PM
To: Frederik Joelving <fcjoelving@gmail.com>

Dear Fredrik,

Difficult to read anything clearly from a gmail. The research is published, and I have nothing else to add.

Almas Heshmati

[Quoted text hidden]

Almas Heshmati <Almas.Heshmati@ju.se>   Sun, Feb 4, 2024 at 10:23 AM
To: Frederik Joelving <fcjoelving@gmail.com>

Dear Frederik,

Hope you are having a nice weekend. A few clarifications to answer your questions:

1. In cases of household and firm level data, one can ignore missing values because of large samples and ability to match and bootstrap. In country level data there is no such option, and the choice is between imputation or losing useful data and its representation.
2. In panel data at country level, we have a few countries each observed several years. Balanced panel data allow to compare countries in each year. Since technology is progressive comparison in same year is important. Imputation allows comparison conditional on same set of countries.
3. The innovation model is complex and adding more explanatory variables increases the frequency of missing units. One-unit missing lead to the whole observation to be lost and none of the remaining full variables can be used in estimation. The sample will not be representing the group of countries like OECD as most less developed countries will be excluded.
4. Countries are different in innovation and data collection which affect the sample, balance of the data, country representation and parameter stability. Thus, missing units cannot be ignored.
5. Trend imputation does not change the result as it is based on same country's data. It is crucial for avoiding yearly gaps in the data and exclusion of countries from the sample because of different missing units in different years. Estimation result based on imputed units of missing and truncated sample are not comparable. The former maintains the sample representation and utilizes fully the non-missing information but in the later case some countries and years are truncated. This makes use of panel data and analysis of dynamics of innovation difficult.
6. Of course, the procedure must be acknowledged and explained. I have missed to explain the imputation procedure in the data section unintentionally in the writing stage of the paper. I am fully responsible for imputations and missing to acknowledge it. Most journals ask for data and codes for replication.
7. When your college approached me and presented himself as a PhD student, I offered him a zoom meeting to explain to him the procedure and even gave him the data. If I had other intensions and did not believe in my imputation approach, I would not share the data with him. If I had to start over again, I would have managed the data in the same way as the alternative would mean dropping several countries and years.
8. One variable was missing the entire period for a country, one can use a matched country data or alternatively use the sample average. This is to not distort the result and avoiding excluding the country from the sample for one variable. This method was used in the past in matching data from neighboring countries with similar characteristics.

Hope I have been to better explain the situation.

[Quoted text hidden]

Frederik Joelving <fcjoelving@gmail.com>   Mon, Feb 5, 2024 at 8:39 AM
To: Almas Heshmati <Almas.Heshmati@ju.se>

Thank you, Almas, I appreciate the time you put into this.

Best,
Fred
[Quoted text hidden]