

Summary:

This study investigates the long-run effects of school starting age (SSA) on educational attainment, labor market outcomes, and health outcome. Using variation from state-level school starting age laws in the U.S., they find that those who are older when entering the school are more likely to report better health and the effect seems to be significant for males only. Self-assessed health is used as main outcome, which is known to be fairly precise reflection of one’s overall health. Authors check the robustness of their estimates by performing falsification test, placebo tests, alternative specifications, as well as adding different set of fixed effects to support that their estimates on long-term health are not influenced by the form of model nor a particular cutoff date. It is very nice to see that authors proactively clarify the identification assumption needed for each estimation model. I have a number of suggestions and questions which would make the paper clearer and stronger.

Major points:

Methodology:

- The average age of peers could be different by school entry law (impact of relative age). Peer effects associated with cohort age composition should be considered.
- I don’t expect the compliance with the law change would be perfect. Parents and teachers can possibly advance or delay school entry for particular children due to various reasons, making some exceptions. What is the fraction of children who were enrolled in kindergarten even if their birthday is later than the cutoff date (i.e. entering school early)? What is the fraction of children who were not enrolled even though they were eligible (i.e. delaying school entry)? How strictly have the cutoff dates been observed? Is there any variation in the compliance rate across the states? Since the compliance rate is not 100%, you should also mention it in your text that your estimate is likely to be overestimated.
- It seems like authors controlled for the race of individuals only in X (a vector of observed characteristics). I would like to know why authors did not control for the demographic characteristics at the time of interview (e.g. household income, marital status, number of children, health insurance status/type, whether receiving any government assistance, etc.), all of which would be highly related to their self-reported health.
- If cutoff changes tend to be bundled with other policies that affect academic and labor market outcomes, it is important to control for these changes in your model. Other education policy controls would include school exit laws, pupil-teacher ratios, teacher’s salaries, and the beginning of state subsidized kindergarten (Bedard and Dhuey, 2012).
- Don’t you also have to control for current state of residence fixed effect? Self-assessed health could be influenced by political, cultural, or environmental factor of states at interview.

Data:

- On page 4: “we restrict our sample to individuals who we can accurately identify their state of residence when they would have started kindergarten” needs more clarification.

- First, what do you mean by “when they would have started kindergarten”? They could have started it any time between the age of 4 and 6, depending on how parents reacted to the minimum school entry age. Are you saying you restrict to individuals whose state of residence can be identified when their age ranges from 4 to 6? Or exactly at 5? From reading the text further, the answer seems to be latter.
- If one seems to have moved during this time period (around the time of kindergarten entry), do you regard him/her as a “mover”? There is no definition about this term in the manuscript.
- You should elaborate more on the format of questions asked in SIPP topical module. Does SIPP ask specifically like “which state were you when you entered the kindergarten?” or are a sequence of questions asked to be able to construct a history of state of residence?
- SIPP migration history topical module is asked only during the wave 2 of each panel. This should be noted in the text.
- I don’t understand how you came up with the formula in footnote 8.
- Due to the limitation of the data, authors are not able to control for any of the respondents’ characteristics at the time of school entry, such as household income, marital status of parents, employment status of parents, family size, etc. Therefore, we cannot make statement that PSSA is random even after controlling for four fixed effects.
- I would focus on outcomes measured after age 30 when educational investments and labor force participation choices are largely complete. Do you choose age 24 as a lower bound in your sample just to be consistent with the previous literature studying long-run effects of SSA?

Finding:

- I am puzzled by the finding that males who are older when they enter school are more likely to report better health, but less likely to complete high school and college. A large body of literature has shown that education and health are positively correlated. I was expecting to see at least the same signs on both outcomes of self-reported health and education level. In addition, authors find that being older at entry has positive effects on male earnings (although not statistically significant), which adds another layer of confusion with the contradictory finding of less likelihood of completing high school. Bedard and Dhuey (2008) show a small and positive impact of backing up of the school start date on educational attainment and positive impact on wages.
- Why the effects are so different between males and females? What is your story behind? Is this what other papers have found as well?
- Why do you partition your sample by gender only? What about partitioning by race/ethnicity?
- Figure 3: Are the plots in Figure 3 showing estimates for both genders pooled? What is the dashed line in the middle? In Tables you have presented the estimates separately by gender. Here you have to clarify that you are presenting the estimate for both genders pooled together, using all cutoff dates as a guideline, so that readers can see to what extent each estimate deviates from this guideline.

Minor Comments:

- Before beginning to use acronym “SSA” for the first time, you must show the acronym when you spelled it out for the last time (second to last paragraph on page 1).
- Need more explanation of a combined effect of school entry and school leaving legislation. Readers who are not familiar with the school system and school starting/leaving legislation don’t understand why students who start school at an older age are more likely to drop out of high school for a longer duration (page 2).
- Page 3 first paragraph: SSA is related to objective health measures, such **as** the number of days.....
- Show data source in Figure 1. Also, add note that the histogram is weighted by observations from your effective sample, and the state entry law dates for 1964-1995. Readers should not have to read the text to make sense of the tables and figures.
- In addition to the histogram of school entry laws (Figure 1), it would be nice if you show a table of school entry law by states and years to provide information on how the law varies across states and how it has changed within states over time.
- In Table 1, you need a table note saying what the numbers in parenthesis mean.
- In Table 3, 4, and 5, show mean of each dependent variable so that readers can have a better sense of the magnitude of each coefficient estimate.

Recommendation to the editor:

This study investigates the effects of school starting age on long-term health, using self-assessed health measure. They use variation in state-level school starting age laws in the U.S. and find that those who are older when entering the school are more likely to report better health and the effect seems to be significant for males only. The effects are robust to the change in functional form, placebo test, and alternative specification.

I have a few concerns, though.

- First, I am not sure how convincing/influential the contribution would be when the paper uses "self-assessed" health as "A" main outcome. Furthermore, they are able to rule out potential channels, but fail to find the exact mechanisms that explain their finding, with just some conjectures.
- Second, I am impressed by how much efforts they have made to test the robustness of their estimates, however, there are many more things that need to be done to improve the paper. I have listed these in the referee report.
- Lastly, I am puzzled by the findings that are contradicting one another -- for example, they find SSA increases self-reported health status but decreases the likelihood of graduating high school. Moreover, the effects are significant only for males, but not for females, which lack explanation in the manuscript.

Therefore, I recommend revise and resubmit (R&R).