The Impact of Family Tax Benefits on Children's Health and Educational Outcomes

Michael Baker, Kourtney Koebel and Mark Stabile

Online Appendix

This appendix provides details on the linkage of the British Columbia Ministry of Education Public School Administrative Data File (BC Data) with the T1 Family File (T1FF). The T1FF data are derived from parent's income tax returns processed by the Canada Revenue Agency (CRA) and contain information on various sources of income including earnings, total income, family benefits, and other income transfer programs. We require the parental tax records for three reasons. First, the tax files provide the net income of all tax filers who are linked to a child in a particular year. Because child benefits in Canada are income-tested (with the exception of the Universal Child Care Benefit (UCCB) which takes the form of a demogrant), we require data on the net income of the highest and lowest earner family member of each child in the education data in order to simulate child benefits for our empirical strategy. Second, we use the tax data to generate control variables that are not available in the education data, including the age of each tax filer linked to the child, household size and number of earners in the household. Finally, to examine whether there are heterogenous impacts of child benefits across the income distribution, we use the T1FF data to assign children to income quintiles computed using average household net income between 2010 and 2013.

For each child in the BC data, we observe every parent T1FF submitted to the CRA between 1994 and 2017. While our empirical strategy only requires income data for tax years between 2010 and 2013, this longer panel of data is beneficial for two reasons. First, it makes it possible to link household income information to each child's education data prior to his/her enrollment in the school system (typically around age 5 in BC). This allows us to retain children who first appear in

the education data post-2013, for whom we would otherwise have no pre-reform income data to use for the simulations. Second, the longer panel of tax data allows us to accurately define the number of children residing in the household in each tax year, as well as their age. This is also required for the simulations since the generosity of child benefits in Canada varies depending on the age and number of the children in the household. More specifically, because the education data is limited to children who are of school age, and does not provide the ability to identify siblings, linking the BC data to the longer panel of parental tax files allows us to, first, tag siblings via parental identification numbers and, second, compute the age of each child in the household in a given year, *so long as the child eventually appears in the BC data*.[1] To derive children's ages, we assign the birth year of each child from the education data to every year of their parent's tax files. We then create variables that capture the age of the youngest, second youngest, third youngest, fourth youngest and fifth youngest child in the household in each year, as required for the CTaCS simulator package.

The linkage of the parental tax data and BC education data is completed by Statistics Canada. For the tax years 2010 onwards, the linkage is based on children's names and birth dates, as parents were only required to provide this information upon claiming child tax credits beginning in 2010. For tax years prior to 2010, Statistics Canada used linkage information from the post-2010 data to retroactively attach children to tax filers. Following the sample restrictions described below, approximately 90 percent of children in the BC education data are successfully linked to at least one tax filer every year; the remaining 10 percent of children for whom no parental tax data exists are excluded from the analysis.

[1] This is particularly important in cases, for example, where an older child enrolled in school has a younger sibling who is not yet enrolled in formal education. If we were to solely use the education data to derive children's ages in each household, our simulations would underestimate child benefits received in the household.

Due to the retroactive linkage process, children might be linked to multiple tax filers, and potentially to multiple households, in a given year. To assign each child to the tax filer and household they are most likely to be residing in, we identify the focal parent (and household) for each child in each year using a series of logistical rules. For children linked to only one tax filer, we assign that filer as the focal parent; approximately 57 percent of child-year observations in the raw T1FF data are linked to one tax filer, while 42 percent of the sample is linked to two tax filers. For children with more than one focal parent, we first select the tax filer linked to the child who claims child tax benefits most frequently; after assigning the most frequent claimant to children, 99.6 percent of the sample is then linked to one main tax filer. For the remaining children who have multiple frequent child benefit claimants, we then assign the female filer as the focal parent. Each child is then assigned to a focal household based on the residency of the focal parent.

Once we have one child observation per year, we then assign income data for tax year *t* to the BC data for the school year also beginning in year *t* (e.g., 2013 tax data is linked to the 2013/2014 school year). Note that while the education data is available until the 2018/2019 school year, the parental tax data ends in 2017. As such, to include 2018 measures of children's health and education outcomes, we assign 2017 household information from the T1FF data to the 2018 education data. This implicitly assumes that the number of earners and household size (i.e., composition) remain constant within a child between 2017 and 2018.

Once the tax and education data are linked, we make the following sample restrictions. First, after computing a measure of average net income between 2010 and 2013 for the highest and lowest earners in the child's focal household, we limit our analysis of mental health and education outcomes to the 2014-2018 school years in order to isolate changes in child benefits stemming from the policy reforms described in our paper. Second, we restrict the sample to children aged 6-

17, born between 1997 and 2008, to ensure that each child has an observation for 2014 and at least one observation in the post-reform period. We also exclude students in grade 13 (i.e., adult learners), children enrolled in kindergarten, any child whose grade is unknown and children for whom net income is missing in 2013. Finally, for the analyses examining educational outcomes, the sample is limited to students who wrote the grade 4 provincial exam prior to the policy reform (in 2014), such that we are comparing the pre-policy grade 4 score to the post-policy grade 7 score. For the educational analyses, we also exclude any children who repeat a grade to avoid differences in test scores over time due to grade retention and further restrict the sample to children who are the appropriate test-writing age given their grade (age 9 in grade 4 and age 12 in grade 7). Finally, because changes in household composition could affect children's health and educational outcomes, we limit our sample to children who are linked to the same tax filers between 2013 and 2017 (as noted above, since the tax data ends in 2017, we assume that the tax filers remain the same in 2018); 84 percent of children in the total sample of children reside in such households.