Online Appendix

1. Issues on identification

a. Addressing the Selective attrition: Inverse Probability Weights

The dataset we use suffers from selective attrition, as many individuals drop out of school, repeat the school year, or change schools. Due to this reason, the remaining students were not representative of the original population and the results may have been affected by attrition bias. The reason is that the individuals who drop out of a panel differ systematically from those who stay in it.

Consider a panel dataset having N individuals surveyed into two different years (t = 1, 2). Let s_{it} denote the selection indicator for each time period, where s_{it} = 1 if both y_{i1} and y_{i2} are observed, and zero when y_{i2} is not observed. Consider (x_{it}, y_{it}) are observed.

Wooldridge (2002) states that "the sequential nature of attrition makes first differencing a natural choice to remove the unobserved effect" (pg. 585):

$$\Delta y_{it} = \beta \Delta x_{it} + \Delta \varepsilon_{it} \qquad t = 2 \qquad (1)$$

Let the score for individual i in the second year be y_{i2}, and in the first year y_{i1}, and let the exogenous variables in the first year be x_{i1}, and in the second year be x_{i2}. Then, y_{i2} is observed only if there is no attrition. With attrition on observables, we can estimate the Inverse Probability Weights (IPW) to solve the problem of sample attrition. This method relies on an auxiliary observed variable (z_{i1}) that needs to be related to the attrition and to the outcome variable (Fitzgerald et al., 1998). The most frequent choice of the auxiliary variable in panel data is a lagged value of y according to Wooldridge (2002) and Fitzgerald et al. (1998). According to Moffit et al. (1999), who also studied sample attrition in panel data, "assuming serial correlation in the y

process, such lagged variables will be related to current values of y conditional on x. If attrition is related to lagged y, least squares projection of y on x using the non-attriting sample will yield biased and inconsistent coefficient estimates. Estimation of attrition probabilities and subsequent weighted least square estimation yields consistent estimation instead" (p. 136). In this study we use the score in Portuguese and Mathematics in 2013 as the z_{i1} variable. It follows that we can write an attrition equation as:

$$s_{it}{}^{*} = \gamma x_{i1} + \delta z_{i1} + v_i \tag{2}$$

We do not observe $s_{it}{}^{*}$, but we do observe s_{it}, which takes the value 1 when both y_{i1} and y_{i2} are observed, and zero when y_{i2} is not observed[1].

Following Wooldridge (2002), ideally, at each t we would observe (y_{it}, x_{it}) for any unit that was in the random sample at t = 1. Instead, we observe (y_{it}, x_{it}) only if s_{it} = 1. According to Wooldridge (2002) "we can easily solve the attrition problem if we assume that, conditional on observables in the first time period, say z_{i1}, (y_{it}, x_{it}) is independent of s_{it}" (p. 587), that is

$$\Pr(s_{it} = 1 | y_{it}, x_{it}, z_{i1}) = \Pr(s_{it} = 1 | z_{i1}) \qquad \text{for t = 2 (or 2017)} \tag{3}$$

The assumption in (3) is called "selection on observables" because we assume that conditional on z_{i1}, selection is independent of (y_{it}, x_{it}) or that the distribution of s_{it} given [z_{i1}, (y_{it}, x_{it})] does not depend on (y_{it}, x_{it}).

There are two steps to obtain the Inverse Probability Weights. First we estimate a probit model of s_{it} on z_{i1} and let $\hat{p}_{it}$ be the fitted probabilities from this model. In the second step the learning score function in year 2 is weighted by $\left(1/\hat{p}_{it}\right)$, while in year 1 the weight is one.

The reasoning behind this procedure is that it gives more weight to individuals that subsequently attrite than to individuals with characteristics that make them more likely to remain in the panel.

b. Addressing the Endogeneity in child labor: the Lewbel Approach[2]

Following Lewbel (2012) and, for simplicity of exposition, simplifying equation (1) reported in the main text, i.e.,

$$y_{it} = \alpha_i + \gamma_t + \delta S_i T + \mu w_{it}^e + \beta x_{it} + \varepsilon_{it}$$

consider the structural equation[3]

$$y = \mu w + \beta_1 x + \varepsilon_1 \tag{4}$$

where

$$w = \beta_2 x + \varepsilon_2 \tag{5}$$

If we have exclusion restrictions, that is, one or more elements of β_1 equal zero and the corresponding elements of β_2 nonzero, we can identify the model using two stage least squares, in which we estimate equation (5) to obtain the fitted values $\widehat{w}$ and then we estimate equation (4) on $\widehat{w}$ and on the subset of *x* that has nonzero coefficients. However, very often we do not have exclusion restrictions and therefore instruments to identify the model. In general, variables affecting *y* also affect *w*.

To circumvent this, the Lewbel identification technique relies on covariates which are correlated with the conditional variance of ε_2 but uncorrelated with the conditional covariance between ε_1 and ε_2. Formally, let *z* be a vector of observed exogenous variables, possibly being a subvector of *x* or even equal to *x*. In this case, Lewbel (2012) shows that under the assumptions

$$cov(z, \varepsilon_1 \varepsilon_2) = 0 \text{ and } cov(z, \varepsilon_2^2) \neq 0 \tag{6}$$

along with heteroskedasticity of ε_2, the structural equation can be identified. In particular, $cov(z, \varepsilon_1 \varepsilon_2) = 0$ assures that the error terms are uncorrelated conditionally to z, and $cov(z, \varepsilon_2^2) \neq 0$ means that z and the variance of the first stage error must be correlated and affects the extent of heteroscedasticity of ε_2. The latter assumption was tested empirically through a modified Wald statistic for groupwise heteroscedasticity in the residuals of a fixed effect regression model, and found that our data satisfy it.[4] If we rewrite the error terms as proposed in Millimet and Roy (2016) as:

$$\varepsilon_1 \equiv \tau + \widetilde{\varepsilon_1}$$

$$\varepsilon_2 \equiv \widetilde{\omega}\tau + \widetilde{\varepsilon_2}$$

where τ is homoscedastic, $\widetilde{\varepsilon_2}$ is heteroskedastic and whose variance depends on z, $\widetilde{\omega}$ are factor loadings, and $\widetilde{\varepsilon_2}$ and $\widetilde{\varepsilon_1}$ are independent of each other and τ, then the conditions in (6) are satisfied. As discussed earlier, in our specific case, τ identifies homoscedastic measurement error in work variables, or an aggregate index of unobserved variables which affect both child labor and test scores, and that is drawn from an identical distribution across observations. The heteroscedastic idiosyncratic component $\widetilde{\varepsilon_2}$ of the child labor can be drawn from different distributions. Typical idiosyncratic shocks to child labor and education performance in Brazil can be represented by adults' unemployment, lack of savings, lack of credit or illness of a household member.

Defining matrices Ψ_{zx} and Ψ_{zz} by

$$\Psi_{zx} = E\left[\begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix}'\right], \ \Psi_{zz} = E\left[\begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix} \begin{pmatrix} x \\ [z - E(z)]\varepsilon_2 \end{pmatrix}'\right]$$

and let Ψ be any positive definite matrix that has the same dimension as Ψ_{zz}, Lewbel shows that,

$$\beta_2 = E(xx')^{-1}E(xw)$$

$$\binom{\beta_1}{\mu} = (\Psi'_{zx}\Psi\Psi_{zx})^{-1}\Psi'_{zx}\Psi\left[\mathrm{E}\binom{x}{[z - E(z)]\varepsilon_2}y\right]$$

This result means that β_2 and μ can be obtained by two stage least squares regression of y on x and w using x and [z - E(z)]ε_2 as instruments. Importantly, the assumption that z is uncorrelated with $\varepsilon_1\varepsilon_2$ means that the generated instrument $[z - E(z)]\varepsilon_2$ is exogenous (since uncorrelated with ε_1) and, so, a valid instrument for w. Also, the larger the degree of heteroskedasticity of ε_2 with respect to z stronger the instrument, since its correlation with w is proportional to the covariance of z and ε_2. The extent of the heteroscedasticity depends on z. According to Lewbel (2018), such a technique is valid also when binary endogenous regressors are used.

The estimation procedure is as follows. The coefficient β_2 is estimated by linearly regressing w on x to obtain the residuals $\hat{\varepsilon}_2$. Then β_1 and μ can be estimated by regressing y on x and w using x and $(z - \bar{z})\hat{\varepsilon}_2$ as instruments, where $\bar{z}$ is the sample mean of z. Let over bars denote sample averages, the resulting estimators are

$$\hat{\beta}_2 = (\overline{xx'})^{-1}\overline{xw}, \quad \hat{\varepsilon}_2 = w - x'\hat{\beta}_2$$

and

$$\binom{\hat{\beta}_1}{\hat{\mu}} = (\hat{\Psi}'_{zx}\hat{\Psi}_{zz}^{-1}\hat{\Psi}_{zx})^{-1}\hat{\Psi}'_{zx}\hat{\Psi}_{zz}^{-1}\left(\frac{\overline{xy}}{(z - \bar{z})\hat{\varepsilon}_2 y}\right)$$

As reviewed and discussed in Fortin and Ragued (2017), in various contexts the results based on the Lewbel approach are found to be more plausible than IV results estimated with external instruments of dubious validity as it is in our case. In theory, when all available instruments are used in the estimation, this should lead to the most asymptotically efficient estimator. For this reason, although we cannot rely on our external instruments alone (i.e., the wage rates for children, men and women as described in footnote 2), we used them together with the Lewbel IV and we found similar results with respect to the estimations where only the Lewbel instruments are used. However, because of some missing values in the external instruments, the estimations with only generated IVs are our preferred estimations. The results of the estimations using both generated and external estimators are available upon request.

2. Construction of the panel dataset when *one* or *two* merging variables are missing

As we said in the main text, once the observations with missing values in *all* merging variables and explanatory variables were dropped, the sample size decreased to 419,562 in 2013 and 262,004 in 2017. At this point, the sample still showed some observations with missing values in *one* or *two* merging variables. In the 2013 sample, the month of birth showed 4,396 missing values and the year of birth showed 3,099. In 2017, among 9th Grade students in complete schools, the month of birth had 692 missing values and the year of birth had 1,177. Due to this, we faced two problems: first, some of our observations with missing values could not be merged and then dropped out of the sample. Secondly, and more importantly, our merging variables were not always able to identify for a 2013 student a unique 2017 correspondent individual (i.e., for such equivocal cases, more than one 2013 individual is associated to one 2017 student).

In order to show the nature of these problems, let us demonstrate by taking one fictive individual in 2013, John (whose name or identification code is unknown from the dataset). According to our merging variables, he was merged with 4 fictive individuals in 2017 (John – his correct pair – Peter, William and Julio). For two of them (John and Peter), all merging variables were not missing, while for William the month of birth was unknown and for Julio the month and year of birth were missing. Our repeated observations (as many times as the number of duplicates – 4 in our example) should then be weighted by W[5], defined as:

$$W_{i,j} = \frac{1}{\sum_j p_{i,j}} p_{i,j}$$

where $p_{i,j}$ is the proportion of non-missing merging variables for 2013's observation *i* (John) and his presumed 2017's pair j. Of course, the sum of weights by *i* should give one (John should indeed be represented by one individual which, in some particular cases as this one, may be the sum of a proportion of different individuals). In our example, we would then have

$W_{i,1} = \frac{1}{3}1 = 0.333$; $W_{i,2} = \frac{1}{3}1 = 0.333$; $W_{i,3} = \frac{1}{3}0.66 = 0.222$;$W_{i,4} = \frac{1}{3}0.33 = 0.111$.

Since there are non-missing values in the school code variable, we considered in the example only 3 merging variables with possible missing observations. This is why we have 0.66 and 0.33, i.e., the proportion of non-missing merging variables for William and Julio, respectively.

It is worth noting that 73% of our final sample had neither missing values in the merging variables, nor duplicated observations. For such cases, the weight is 1.

3. Estimation of the Inverse Probability Weights (IPW)

The results reported in Table OA1 show the coefficients used to estimate the IPW. It is worth remarking that the coefficients of the score variables were highly statistically significant, indicating that attrition bias might be present when estimating children´s school performance models.[6] Also, the larger the children's test scores, the larger the probability of staying in school, and therefore in the sample, as we expected.

The final weight used in the descriptive and econometric estimations is then the product of the inverse probability weight and the weight defined in section 2 above. As mentioned earlier, the inverse probability weight in 2013 is 1.

4. Additional description statistics on child labor

From table OA2, in 5th Grade, close to 55% of girls and 63% of boys worked neither in the household, nor in the labor market. Girls worked more in the household (39%), compared to boys (23%). On the other hand, 10% of boys worked only in the labor market compared to 3% of girls. The percentage of boys working in both the household and the labor market (9%) was also larger than girls (3%). The average hours spent on household chores per day was larger when the children worked in both household chores and labor market, spending around 2.7 hours per day. When they worked in the household only, girls and boys spent approximately 2.5 hours per day. The percentage of students working increased with age (or Grade): for 9th Grade students, 10% of boys worked in the labor market only and 7% in both the household and labor market. The largest increase for older girls concerned household chores (from 39% in the 5th Grade to 61% in the 9th grade).

The number of hours a child has spent working in his/her own household per day is presented in Table OA3. It should be noticed that girls, not only worked more in the household than boys, but they also spent more hours doing household tasks. In the 9th Grade, 9.3% of girls worked more than 3 hours a day in household activities, compared to 4.5% of boys. Some studies show that giving children household chores helps to form accountability and self-confidence and that they are more likely to succeed in adulthood (Rossmann, 2002). However, if a child is overloaded with household chores, working a large number of hours per day can harm his or her future life as less time is allocated to studying and doing homework. Due to this reason, when a child claimed to be performing household tasks for one hour or less per day, we considered that he/she was not working. From our data, about 23.8% of girls in 9^{th} Grade spent 2 or more hours a day on household chores.

Additional Tables

Table OA1 – Coefficients of the Probit Model.

Variables	Coefficients of the Probit Model Dependent variable is equal to 1 if student is in 2013 and in 2017 and 0 otherwise	
Scores_portuguese	0.246***	-
	(0.0016)	-
Scores_mathematics	-	0.249***
	-	(0.0017)
	-0.960***	-0. 961***
Constant	(0.0017)	(0.0017)
Pseudo R^2	3.01%	3.02%
Observations	775,554	775,554

Source: Authors' estimation based on Microdata of Prova Brasil 2013 and 2017

Table OA2 - Number and percentage of 5^{th} and 9^{th} Grade students, according to their work status,° by gender

	2013 - 5^{th} Grade			2017 – 9^{th} Grade		
Work Status	number	%	Average hours/day spent in household chores	number	%	Average hours/day spent in household chores
Girls						
Do not work*	28,374	54.91	0. 77	16,088	31.13	0.84
Work only in the hh	20,035	38.77	2.48 (14.7)	31,279	60.53	2.48 (13.4)
Work only in the market	1,554	3.01	0.82	1,322	2.56	0.81
Work in both	1,715	3.32	2.68 (23.8)	2,989	5.78	2.64 (20.3)
Total	51,678		1.50 (6.5)	51,678	100	1.94 (9.3)
Boys						
Do not work*	26,911	62.76	0.63	21,591	50.35	0. 70
Work only in the hh	9,706	22.64	2.45 (14.6)	13,978	32.6	2.35 (10.4)
Work only in the market	3,917	9.14	0.74	4,339	10.12	0.64
Work in both	2,345	5.47	2.64 (22.6)	2,971	6.93	2.47 (16.4)
Total	42,879	100	1.16 (4.5)	42,879	100	1.34 (4.5)

Source: Authors' estimation based on Microdata of Prova Brasil 2013 and 2017.
*Considered not working if worked 1 hour or less in the household per day.
º Numbers in parentheses show the share of observations spending 4 hours or more per week in household chores.

Table OA3 - Number and percentage of 5^{th} and 9^{th} Grade students, according to the number of hours per day they worked in their household by gender.

Hours working hh/day	2013 – 5th Grade				2017 – 9th Grade			
	Girls		Boys		Girls		Boys	
	number	%	number	%	number	%	number	%
not work in hh	6,763	13.09	10,953	25.54	2,883	5.58	8,706	20.3
less than 1 hr/day	23,165	44.83	19,875	46.35	14,527	28.11	17,224	40.17
From 1 to 2 h/day	14,359	27.79	8,131	18.96	21,996	42.56	12,614	29.42
From 2 to 3 h/day	4,038	7.81	1,976	4.61	7,483	14.48	2,393	5.58
More than 3 hr/day	3,353	6.49	1,944	4.53	4,789	9.27	1,942	4.53
Total	51,678	100	42,879	100	51,678	100	42,879	100

Source: Authors' estimation based on Microdata of Prova Brasil 2013 and 2017.

Table OA4 - The students´ level of performance in the Portuguese and Mathematics test scores, according to their scores.

Level	5th Grade				9th Grade			
	Portuguese		Mathematics		Portuguese		Mathematics	
	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
level 0	0	125	0	125	0	125	0	200
level 1	125	150	125	150	125	150	200	225
level 2	150	175	150	175	150	175	225	250
level 3	175	200	175	200	175	200	250	275
level 4	200	225	200	225	200	225	275	300
level 5	225	250	225	250	225	250	300	325
level 6	250	275	250	275	250	275	325	350
level 7	275	300	275	300	275	300	350	375
level 8	300	325	300	325	300	325	375	400
level 9	325	350	325	350	325	350	400	425
level 10	-	-	-	-	-	-	350	375
level 11	-	-	-	-	-	-	375	400
level 12	-	-	-	-	-	-	400	425

Source: Prova Brasil 2013 and 2017.

Table OA5 – Coefficients of the fixed-effect models, full specification with IPW and IV for test scores in Portuguese and Mathematics, by gender

Variables	Portuguese		Mathematics	
	girls	boys	girls	boys
	(5)	(10)	(5)	(10)
Portuguese test Score	-0.076***	-0.086***	-0.077***	-0.147***
Mathematics test Score	-0.464***	-0.308***	-0.338***	-0.311***
don't work	-0.557***	-0.517***	-0.494***	-0.488***
Work only in the hh	0.001	-0.005	-0.126***	0.030*
Work only in the market	0.082*	0.105**	0.015	0.022
Work in both	0.152***	0.145***	0.107**	0.087**
1 if repeat school year	0.163***	0.112**	0.120**	0.088*
Number of people in hh	-0.042***	-0.018***	-0.025***	-0.005
Number of cars in hh	0.266***	0.227***	0.251***	0.208***
1 if floor in school	0.328***	0.317***	0.307***	0.305***
1 if child starts 2 to 4 years	0.179***	0.177***	0.155***	0.177***
1 if child starts 4 to 6 years	0.001	-0.014	-0.003	-0.001
1 if starts school at 6 or 7	0.004	-0.008*	-0.003	-0.004
Age of the teacher	-0.000	0.000	0.000	0.000

Source: Source: Authors' estimation based on Microdata of Prova Brasil 2013 and 2017.
Notes: **significant at 1% level, ** at 5% level, * at 10% level

Table OA6 – Coefficients of the fixed-effect models with IPW and IV for test scores in Portuguese and Mathematics in 2007/2011; 2009/2013; 2011/2015 and 2013/2017 panels, for girls and boys.

Variables	Panel A: Portuguese							
	Girls				Boys			
	2007-2011	2009-2013	2011-2015	2013-2017	2007-2011	2009-2013	2011-2015	2013-2017
Work only at home	-0.183***	-0.160***	-0.211***	-0.076***	-0.210***	-0.208***	-0.242***	-0.086***
	(0.016)	(0.015)	(0.014)	(0.016)	(0.017)	(0.018)	(0.019)	(0.022)
Work only in the market	-0.668***	-0.306***	-0.443***	-0.464***	-0.535***	-0.287***	-0.281***	-0.308***
	(0.063)	(0.040)	(0.043)	(0.043)	(0.061)	(0.029)	(0.030)	(0.032)
Work in both	-0.449***	-0.420***	-0.706***	-0.557***	-0.477***	-0.362***	-0.445***	-0.517***
	(0.071)	(0.058)	(0.066)	(0.050)	(0.033)	(0.032)	(0.036)	(0.038)
R-squared	0.010	0.019	0.009	0.021	0.019	0.035	0.018	0.029
underidentification test (Kleibergen-Paap rk LM statistic)	325.09***	435.00***	396.88***	571.77***	637.92***	2272.67***	3302.76***	5496.81***
weak identification test (Kleibergen-Paap Wald F statistic) °	17.75	40.06	28.12	89.79 (19.94)	38.02	281.15	304.29	461.78 (19.67)
Hansen J statistic (overidentification test)	57.292	52.853	49.975	22.84	75.135	25.619	29.205	25.25
	Panel B: Mathematics							
Work only at home	-0.138***	-0.131***	-0.171***	-0.077***	-0.197***	-0.191***	-0.256***	-0.147***
	(0.015)	(0.014)	(0.014)	(0.015)	(0.018)	(0.019)	(0.019)	(0.022)
Work only in the market	-0.516***	-0.277***	-0.321***	-0.338***	-0.440***	-0.262***	-0.321***	-0.311***
	(0.057)	(0.038)	(0.039)	(0.041)	(0.057)	(0.030)	(0.030)	(0.032)
Work in both	-0.380***	-0.399***	-0.522***	-0.494***	-0.409***	-0.408***	-0.486***	-0.488***
	(0.066)	(0.055)	(0.059)	(0.050)	(0.033)	(0.033)	(0.036)	(0.040)
R-squared	0.012	0.020	0.015	0.023	0.014	0.021	0.016	0.027
underidentification test (Kleibergen-Paap LM statistic)	367.22***	472.10***	436.86***	588.39***	737.95***	2377.37***	4203.58***	5485.42***
weak identification test (Kleibergen-Paap Wald F statistic) °	20.44	45.55	32.46	96.79 (19.94)	21.74	139.80	171.58	231.07 (20.59)
Hansen J statistic (overidentification test)	35.131	29.801	43.417	29.29	122.158	43.867	69.489	35.33
States x Trend	yes	yes	yes	yes	yes	yes	yes	yes
Individual fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
Year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
Exogenous variables#	yes	yes	yes	yes	yes	yes	yes	yes
Observations	248,800	146,068	180,464	103,356	206,984	131,294	158,452	85,758

Source: Authors' estimation based on Microdata of Prova Brasil.
Note: ***, **, * significant at 1% level, 5% and 10% level respectively.
#All columns have the control variables shown in table OA5.
° The value in brackets reported in the row with the weak identification test indicates the Stock-Yogo critical value at 5%.

Table OA7 – Coefficients of the fixed-effect models with IPW and IV for test scores in Portuguese and Mathematics in the 2013/ 2017 panel, girls and boys, using full sample.

Variables	***Panel A: Portuguese***	
	Girls	***boys***
	With IPW and With IV	With IPW and With IV
	full	full
Work only at home	-0.080***	-0.114***
	(0.015)	(0.021)
Work only in the market	-0.483***	-0.310***
	(0.040)	(0.030)
Work in both	-0.636***	-0.583***
	(0.053)	(0.036)
R-squared	0.012	0.020
	Panel B: Mathematics	
Work only at home	-0.082***	-0.169***
	(0.014)	(0.021)
Work only in the market	-0.364***	-0.324***
	(0.036)	(0.029)
Work in both	-0.569***	-0.556***
	(0.051)	(0.037)
R-squared	0.013	0.018
States x Trend	yes	yes
Individual fixed effect	yes	yes
Year fixed effect	yes	yes
Exogenous variables#	yes	yes
Observations	140,350	118,726

Source: Authors' estimation based on Microdata of Prova Brasil 2013 and 2017.
Note: The sample used here also includes those observations having 1 or 2 missing values in the merging variables, as explained in section 2 above. ***, **, * significant at 1% level, 5% and 10% level respectively. We also included the variables in table OA5.

Notes

[1] Note that, for simplicity, in this subsection x_{it} identify all explanatory variables, including child labor.
[2] For simplicity of the exposition, we did not indicate the time dimension in the presentation of this approach.
[3] For simplicity, we present the logic with just one endogenous explanatory variable, but the case with multiple endogenous regressors can be easily extended.
[4] The test was run through the Stata command "xttest3". The results of the test can be obtained upon request.
[5] Such approach takes some inspiration from the probabilistic linkage literature. One example is Ridder and Moffit (2005).
[6] This is also confirmed by the BGLW attrition test (available upon request) according to which the hypothesis of equality of coefficients estimated on the full and the non-attriting samples is strongly rejected.

References

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*: Journal of the Econometric Society, 1287-1294.

Fitzgerald, J., Gottschalk, P., & Moffitt, R. (1998). An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics. *The Journal of Human Resources, 33*(2), 251- 299.

Fortin, B., & Ragued, S. (2017). Does temporary interruption in postsecondary education induce a wage penalty? Evidence from Canada. *Economics of Education Review*, *58*, 108-122.

Millimet, D. L. & Roy, J. (2016). Empirical tests of the pollution haven hypothesis when environmental regulation is endogenous. *Journal of Applied Econometrics*, 31 (4), 652–677.

Moffit, R., Fitzgeral J. & Gottschalk, P. (1999). Sample Attrition in Panel Data: The role of selection on observables. Annales d'Économie et de Statistique, 55-56, 129-152.

Ridder, G., & Moffitt, R. (2007). The econometrics of data combination. *Handbook of econometrics*, *6*, 5469-5547.

Rossmann, M. (2002). Involving Children in Household Tasks: Is it worth the Effort? ResearchWorks, University of Minnesota, *mimeo*.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.