Early Child Development and Care, 2015
Vol. 185, No. 2, 291–316, http://dx.doi.org/10.1080/03004430.2014.913586

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

Long-term benefits of full-day kindergarten: a longitudinal population-based study

M.D. Brownell[a,b]*, N.C. Nickel[a,b], D. Chateau[a,b], P.J. Martens[a,b], C. Taylor[b], L. Crockett[a,b], A. Katz[a,b], J. Sarkar[b], E. Burland[b], C.Y. Goh[b] and the PATHS Equity Team

[a]*Department of Community Health Sciences, University of Manitoba, Manitoba, Canada;* [b]*Manitoba Centre for Health Policy, University of Manitoba, Manitoba, Canada*

(*Received 5 March 2014; final version received 7 April 2014*)

In the first longitudinal, population-based study of full-day kindergarten (FDK) outcomes beyond primary school in Canada, we used linked administrative data to follow 15 kindergarten cohorts (*n* ranging from 112 to 736) up to grade 9. Provincial assessments conducted in grades 3, 7, and 8 and course marks and credits earned in grade 9 were compared between FDK and half-day kindergarten (HDK) students in both targeted and universal FDK programmes. Propensity score matched cohort and stepped-wedge designs allowed for stronger causal inferences than previous research on FDK. We found limited long-term benefits of FDK, specific to the type of programme, outcomes examined, and subpopulations. FDK programmes targeted at low-income areas showed long-term improvements in numeracy for lower income girls. Our results suggest that expectations for wide-ranging long-term academic benefits of FDK are unwarranted.

Keywords: full-day kindergarten; population-based; propensity score; stepped wedge; socioeconomic status; academic achievement; assessment

The popularity of full-day kindergarten (FDK) has grown over recent decades, largely due to changing family dynamics and dual parental workforce participation as well as the increasing recognition for the need to improve childhood equity. Currently in Canada, FDK provision and accessibility varies greatly across the country, with some provinces having established or moving towards full implementation, and other provinces offering limited access to FDK programmes (Preston, Cottrell, Pelletier, & Pearce, 2012). With the recent move to province-wide FDK by Canada's most populous province, Ontario (population = 13.5 million), the value of FDK has been fiercely debated across the country. FDK proponents contend the programme is a means of improving children's school readiness and suggest that experiences and performance during this period set the stage for future academic success (Pascal, 2009). Opponents point to research claiming that there are potential negative impacts of FDK (Cooper,

*Corresponding author. Email: marni_brownell@cpe.umanitoba.ca

© 2014 The Author(s). Published by Taylor & Francis.
This is an Open Access article. Non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly attributed, cited, and is not altered, transformed, or built upon in any way, is permitted. The moral rights of the named author(s) have been asserted.

Allen, Patall, & Dent, 2010) and that the long-term cost–benefit ratio is expected to be close to zero (DeCicca, 2007; Reynolds & Temple, 2008).

The potential positive benefits of FDK include higher academic achievement, fewer grade retentions, easier transitions into first grade, better socialisation and self-esteem, less hurried instruction, fewer transitions, lower child care costs, more learning opportunities for low-income children, and greater academic equity (Cooper et al., 2010; Hahn et al., 2014). On the other hand, there are several potentially negative impacts of FDK, including higher expectations, pushing first-grade material down to kindergarten, increased child fatigue, separation anxiety, less time for informal learning, less planning time for teachers, still unmet childcare needs, diminished parent responsibilities, increased costs, continued unequal access to FDK, and that it may take resources away from more effective interventions (Cooper et al., 2010).

Research suggests that FDK benefits children both academically and socially in the short term, and there is increasing evidence that FDK is especially beneficial for children of low socioeconomic status (SES) or educationally disadvantaged backgrounds (Bingham & Hall-Kenton, 2012; Chang, 2012; Chang & Singh, 2008; Preston et al., 2012; Puleo, 1988; Schroeder, 2007). Studies have reported that FDK increases literacy skills compared to half-day kindergarten (HDK) (Elicker & Mathur, 1997; Walston & West, 2004). Similar to the meta-analysis results by Cooper et al. (2010), studies derived from the Early Childhood Longitudinal Study – Kindergarten (ECLS-K) data generally demonstrated a greater positive gain in reading and math for children in FDK at the end of kindergarten and heading into grade one, and consistently supported the short-term benefits of FDK (Lee, Burkam, Ready, Honigman, & Meisels, 2006; Yan & Lin, 2005).

Although FDK seems to present some benefits early on, it has been noted that these gains fade over time (Cannon, Jacknowitz, & Painter, 2006; Cooper et al., 2010; Olsen & Zigler, 1989; Votruba-Drzal, Li-Grining, & Maldonado-Carreno, 2008; Wolgemuth, Cobb, Winokur, Leech, & Ellerby, 2006). In a meta-analysis of FDK, Cooper et al. (2010) found that the positive effects on achievement seen at the end of the kindergarten year completely diminished by the end of third grade. Taking a developmental approach using individual academic trajectories, Votruba-Drzal et al. (2008) used the ECLS-K cohort data and discovered that this fade out occurs due to steeper academic trajectory growth from the spring of kindergarten to the spring of fifth grade among HDK students. Votruba-Drzal et al. attributed the faster HDK growth trajectory to cognitive stimulation in the home environment and family poverty, suggesting that academic trajectories and school performance may in part be explained by lower levels of poverty and more stimulation in the home environment among those attending HDK as opposed to those in FDK.

The different socioeconomic characteristics of families attending FDK and HDK programmes (Votruba-Drzal et al., 2008) call into question the validity of many FDK evaluations, and there is a need for study designs that allow for ‘strong causal inferences’ (Cooper et al., 2010, p. 34), as well as for longer-term follow-up of FDK students (Reynolds & Temple, 2008). The current study addressed these gaps, using provincial education data from the Canadian province of Manitoba. The objective of this study was to assess whether FDK, when compared with HDK, was associated with improved longer-term academic performance. Using linked administrative data, we were able to follow 15 kindergarten cohorts up to grade 9, analysing provincial assessments conducted in grades 3, 7, and 8 and course marks and credits earned in grade 9 to determine achievement. This is the first longitudinal, population-based study of FDK outcomes beyond primary school in Canada.

Methods

Data sources

Data came from the Population Health Research Data Repository held at the Manitoba Centre for Health Policy (MCHP), University of Manitoba. The Repository contains a collection of population-based data on health, education and social services used by residents of the province of Manitoba (population = 1,265,000), linkable at the individual level using an encrypted identifier. The specific datasets used for this study included: enrolment, marks, and assessment data for all Manitoba school children from the provincial department of Education; the Manitoba population registry, updated every six months, to determine demographic information on students (e.g. area of residence, number of residential moves, family information); provincial social assistance data to identify students living in poverty; Canada census data to determine area-level SES; and child welfare data to identify families involved in child protection services. The data in the Repository have been used extensively for research (Brownell et al., 2002, 2006, 2010; Oreopoulos, Stabile, Walld, & Roos, 2008; Roos, Gupta, Soodeen, & Jebamani, 2005; Roos & Nicol, 1999).

Population

Education enrolment information from three separate school divisions in Manitoba (labelled 1, 2, and 3 here) was used to identify nine cohorts of children enrolled in kindergarten in school divisions 1 and 2 from school years 1998/1999 to 2003/2004 and 2006/2007 to 2007/2008, and six cohorts of children enrolled in kindergarten in school division 3, from 1998/1999 to 2003/2004.

Exposure

In school division 1, FDK was introduced gradually to the schools with the lowest area-level SES, as measured using Canada census data, starting in 1998/1999, with all five target schools included by 2001/2002. In school division 2, only HDK operated throughout the study period; school division 2 was used to identify control cohorts for comparison with school division 1 (see Table 1 for cohort numbers before and after matching; further information on matching can be found in section 'Study Design'). In school division 3, a decision was made to implement FDK in all schools with kindergarten classes; however, it was introduced gradually, starting with three schools in 1999/2000 and including all 20 schools by 2003/2004 (see Table 2 for cohort numbers; further information on stepped-wedge design can be found in section on 'Study Design').

Kindergarten is not mandatory in Manitoba, however, it is universally available and most children enroll (98% of eligible children enrolled in 2012/2013). A few school divisions offer FDK in all schools, some offer FDK in targeted schools, and most offer HDK. The focus in kindergarten is on play-based programmes with classroom programming varying according to local needs, such as English as an additional language learners, multi-age classrooms, and special needs students. Development of skills in literacy and communication, problem-solving, cooperation, and technology use are emphasised. Some rural kindergarten programmes actually run for the full school day (approximately 5.5 hours of instructional time) but on alternate days, for efficient use of bus services. The school year in Manitoba runs from the beginning

Table 1. Kindergarten cohort size before and after matching, by each outcome measured, school divisions 1 and 2.

Outcome	Grade 3 reading and numeracy				Grade 7 student engagement and math				Grade 8 reading and writing				Grade 9 achievement index			
	Before matching		After matching		Before matching		After matching		Before matching		After matching		Before matching		After matching	
Kindergarten cohort	FDK	HDK	FDK	HDK	FDK	HDK	FDK	HDK	FDK	HDK	FDK	HDK	FDK	HDK	FDK	HDK
1998/1999													21	1090	21	91
1999/2000									63	980	63	306	71	1089	71	323
2000/2001					65	488	62	176	66	991	66	307	75	1107	74	332
2001/2002					50	428	50	167	48	891	48	259	58	997	58	314
2002/2003					168	405	165	346	169	740	164	572				
2003/2004					143	409	137	332	139	729	134	479				
2004/2005					133	382	130	276								
2007/2008	163	758	159	425												
2008/2009	130	779	125	444												

Table 2. Kindergarten cohort size before and after implementation of FDK, school division 3.

Outcomes	Grade 9	Grades 8 and 9	Grades 7–9	Grade 7–9	Grades 7 and 8
Kindergarten cohort	Three schools	Six schools	Six schools	two schools	Three schools
1998/1999	66	138	66	s	69
1999/2000	**53**	131	67	s	43
2000/2001	**66**	**120**	60	0	64
2001/2002	**72**	**131**	**69**	s	47
2002/2003	**62**	**141**	**53**	**11**	58
2003/2004	**70**	**122**	**88**	**10**	**41**

Values that are not in bold represent students in half-day kindergarten before implementation of FDK; bolded values represent students in FDK. 's' indicates cell sizes suppressed due to low count.

of September to the end of June, approximately 185 days of instruction. FDK programmes are effectively twice the amount of instructional time as HDK programmes.

Outcomes

Six outcomes were examined in the comparison using students from school divisions 1 and 2: grade 3 reading and numeracy assessments; grade 7 mathematics and student engagement assessments; grade 8 reading comprehension/writing assessment; and a grade 9 achievement index, which combined student-level information on number of credits earned and marks for coursework (Roos et al., 2013). The assessments in grades 3 through 8 were teacher administered and determined whether students were meeting expectations for grade. Teachers are provided with documentation as well as professional development sessions regarding these assessments, and the results are entered into the provincial database at the student level using a web-based application. See Table A1 for a more detailed description of the outcomes. Four outcomes were examined for the students from school division 3: grades 7 (mathematics and engagement) and 8 (reading comprehension/writing) assessments and grade 9 achievement index.

Study design

Analytic framework

Our study design followed Rubin's Potential Outcomes Framework (Rubin, 2005, 2008). In this framework, each child is considered to have two sets of potential outcomes: (1) when she/he received FDK and (2) when she/he received HDK. The difference in these two sets of outcomes is the impact of FDK compared with HDK. The fundamental challenge with such research is that, for each child, we can only observe one set of outcomes – if the child received FDK, we cannot know her academic outcomes with HDK and vice versa (Holland, 1986).

Our task was to design a study that provides unbiased estimates of these potential outcomes. A randomised study will yield unbiased estimates of the average effect of FDK, compared with HDK, on outcomes. When random assignment is unfeasible, observational studies often utilise techniques such as multiple regression; however, multiple regression often fails to fully adjust for differences in observed covariates

between exposure groups (Rubin, 2001). We therefore utilised two alternative observational study designs to identify differences in outcomes attributable to FDK *vs.* HDK: (1) Propensity Score Matched Cohort Design and (2) Stepped-Wedge Design. Both study designs are described below.

(1) *Propensity Score Matched Cohort Design.* Analyses comparing school divisions 1 and 2 followed a matched cohort design where children enrolled in FDK were matched to children in HDK, based on their estimated propensity score, and followed to grade 9 to assess differences in outcomes. A child's propensity score is simply her/his conditional probability of being enrolled in FDK (compared with HDK), given her/his observed covariates (Rubin, 2001; Rubin & Thomas, 1996; Guo & Fraser, 2009; D'Agostino, 1998). The propensity score is useful for (a) balancing differences in observed covariates between FDK and HDK children and, subsequently, (b) creating comparable exposure groups (D'Agostino, 1998; Guo & Fraser, 2009; Morgan & Winship, 2007; Rubin, 2001, 2008). Furthermore, conditional on the propensity score, observed covariates are independent of assignment to FDK *vs.* HDK (D'Agostino, 1998; Guo & Fraser, 2009; Morgan & Winship, 2007; Rubin, 2001, 2008). Thus, assessing and comparing outcomes between children receiving FDK and children receiving HDK with the same value on the propensity score, follows Rubin's potential outcomes framework and facilitates estimation of the differences in performance attributable to FDK.

Each child's propensity score was estimated by logistic regression. Several variables were used to predict exposure to FDK *vs.* HDK that have been shown to affect school performance (i.e. they were true confounders of the relationship between FDK and school performance): an area-level index of SES derived from the Canada census at the dissemination area (DA) level (approximately 400 people) that includes average household income, employment, education, and lone-parent status (Brownell et al., 2006; Chateau, Metge, Prior, & Soodeen, 2012; Guèvremont, Roos, & Brownell, 2005); total number of residential moves (Cutuli et al., 2013; Ingersoll, Scamman, & Eckerling, 1989; Scanlon & Devine, 2001); mother's age at first birth (Jaffee, 2002; Jutte et al., 2010); mother's marital status at kindergarten start date (Strohschein, Roos, & Brownell, 2009); family income assistance receipt (Chase-Lansdale, Coley, Lohman, & Pittman, 2002; Roos et al., 2006); family involvement with child welfare services (Brownell et al., 2010; Geenen & Powers, 2006; Scherr, 2007); child's sex; and child's age in months. For grade 3 assessments, a school readiness measure at school entry, based upon the Early Development Instrument (EDI; Forget-Dubois et al., 2007; Janus & Offord, 2007) was also used. The logistic regressions were used to generate predicted probabilities of exposure to FDK (*vs.* HDK) for each child, based on his/her observed covariates. These predicted probabilities were used as the estimated propensity scores.

The estimated propensity scores were then used to match each FDK child with up to five HDK children. Specifically, we used greedy, balanced, nearest neighbour matching within a defined calliper of up to 0.05 (Guo & Fraser, 2009). After matching FDK to HDK children, based on the propensity score, we assessed balance on observed covariates to determine whether any differences between FDK and HDK children remained. Absence of any differences on observed covariates suggests comparability between the two groups.

(2) *Stepped-wedge design.* In a stepped-wedge design, the intervention is implemented sequentially over a particular timeframe, to either individuals or groups of individuals. By the end of the timeframe, all individuals or groups will have received or be receiving the intervention. The ordering of the intervention is not based on

participant characteristics, but is determined randomly. The individuals or groups in the study act as their own controls, providing data points in both the non-intervention and in the intervention group (Brown & Lilford, 2006). In school division 3, the FDK programme could not be implemented in all 20 schools at once for practical reasons and was introduced gradually with three to six new schools starting FDK each year over the course of five years (Table 2). The ordering of schools for beginning FDK was not based on socioeconomic or school achievement factors; however, those schools where more parents were in favour of FDK may have been the first selected, which could reflect maternal employment factors.

Outcome analyses

For school divisions 1 and 2, the propensity score matched cohort was used to compare differences in academic performance between FDK and HDK children. Rates for each of the six outcomes were calculated by income quintile groupings derived from the 2006 Canada census. Average household income was assigned at the DA level, and DAs were grouped into quintiles to represent approximately 20% of the Manitoba population, with Q1 being the lowest and Q5 being the highest (Brownell et al., 2002). Each outcome was generated for the FDK children and for their matched HDK controls in that quintile. There were fewer children in higher income quintiles, so depending on the outcome, the higher income quintiles were grouped together.

For school division 3, a stepped-wedge design was used where each school acted as its own control. Income quintiles were generated as described above for school divisions 1 and 2; however, in school division 3, there were fewer children in the lower income quintiles so when necessary the three lowest income quintiles were grouped.

For both analyses (school division 1 compared to 2, and school division 3 compared to itself in earlier years), logistic regression using generalised linear mixed models was run, with the school and academic year at the time of outcome as a clustering level to control for potential school effects in a given year. For school divisions 1 and 2, income quintile was entered as a covariate in the models. For school division 3, income quintile, age, sex, child welfare involvement, income assistance receipt, age of mother at first birth, number of moves, mother's marital status, age, and sex were entered as covariates in the models. To determine whether FDK had a differential effect according to income quintile, a kindergarten type (FDK or HDK) by income quintile interaction was entered into the models for all analyses. Based on previous findings (Heckman, Pinto, & Savelyev, 2013), an interaction between kindergarten type and sex was also modelled.

Drawing inferences about the plausibility that FDK has an effect on school performance rests on the strong assumption that we have controlled for all measured and unmeasured confounding. Although it is impossible to know the strength of any residual confounding, we assessed how strong such confounding would need to be in order to invalidate statistically significant findings, using gamma sensitivity analyses outlined by Rosenbaum (Guo & Fraser, 2009; Jiang & Foster, 2013; Rosenbaum, 2010). These analyses provide an assessment of the plausibility that observed associations may reflect underlying causal relationships.

Results

Table 3 shows the demographic information for students in school divisions 1 and 2 for the matched sample (Table A2 shows this information for the cohorts before and after

Table 3. Demographic characteristics of children in FDK and HDK, by outcomes examined, for matched sample in school divisions 1 and 2.

		Grade 3			Grade 7			Grade 8			Grade 9		
Variables		FDK count (%)	HDK count (%)	*p*-value	FDK count (%)	HDK count (%)	*p*-value	FDK count (%)	HDK count (%)	*p*-value	FDK count (%)	HDK count (%)	*p*-value
Number of moves (years)	≤2	206 (72.54)	658 (75.72)		406 (74.63)	1000 (77.10)		347 (73.05)	1409 (73.27)		148 (66.07)	722 (68.11)	
	>2	78 (27.46)	211 (24.28)	0.28	138 (25.37)	297 (22.90)	0.26	128 (26.95)	514 (26.73)	0.92	76 (33.93)	338 (31.89)	0.55
Sex	Male	158 (55.63)	468 (53.86)		306 (56.25)	703 (54.20)		259 (54.53)	1003 (52.16)		112 (50.00)	554 (52.26)	
	Female	126 (44.37)	401 (46.14)	0.6	238 (43.75)	594 (45.80)	0.42	216 (45.47)	920 (47.84)	0.35	112 (50.00)	506 (47.74)	0.54
Mother's marital status (at kindergarten start date)	Married	98 (34.51)	342 (39.36)		273 (50.18)	692 (53.35)		235 (49.47)	982 (51.07)		99 (44.20)	483 (45.57)	
	Not married	186 (65.49)	527 (60.64)	0.14	271 (49.82)	605 (46.65)	0.21	240 (50.53)	941 (48.93)	0.53	125 (55.80)	577 (54.43)	0.71
Income quintile	Q1 (lowest)	82 (28.87)	140 (16.11)		120 (22.06)	243 (18.74)		102 (21.47)	331 (17.21)		64 (28.57)	249 (23.49)	
	Q2	88 (30.99)	249 (28.65)		198 (36.40)	363 (27.99)		190 (40.00)	550 (28.60)		116 (51.79)	367 (34.62)	
	Q3	68 (23.94)	239 (27.50)		127 (23.35)	254 (19.58)		105 (22.11)	463 (24.08)		31 (13.84)	252 (23.77)	
	Q4	33 (11.62)	203 (23.36)		67 (12.32)	299 (23.05)		53 (11.16)	429 (22.31)		9 (4.02)	150 (14.15)	
	Q5 (highest)	13 (4.58)	38 (4.37)	<0.0001	32 (5.88)	138 (10.64)	<0.0001	25 (5.26)	150 (7.80)	<0.0001	4 (1.79)	42 (3.96)	<0.0001
Income assistance	Yes	119 (41.90)	259 (29.80)		196 (36.03)	381 (29.38)		182 (38.32)	627 (32.61)		98 (43.75)	437 (41.23)	
	No	165 (58.10)	610 (70.20)	<0.001	348 (63.97)	916 (70.62)	<0.01	293 (61.68)	1296 (67.39)	<0.05	126 (56.25)	623 (58.77)	0.49
Children in care/ protective services	Yes	59 (20.77)	139 (16.00)		116 (21.32)	227 (17.50)		98 (20.63)	340 (17.68)		41 (18.30)	194 (18.30)	
	No	225 (79.23)	730 (84.00)	0.06	428 (78.68)	1070 (82.50)	0.05	377 (79.37)	1583 (82.32)	0.14	183 (81.70)	866 (81.70)	1

EDI domain – not ready in physical health	Yes	53 (18.66)	90 (10.36)										
	No	231 (81.34)	779 (89.64)	<0.001									
EDI domain – not ready in emotional maturity	Yes	45 (15.85)	96 (11.05)										
	No	239 (84.15)	773 (88.95)	<0.05									
EDI domain – not ready in social competence	Yes	36 (12.68)	72 (8.29)										
	No	248 (87.32)	797 (91.71)	<0.05									
EDI domain – not ready in language development	Yes	23 (8.10)	53 (6.10)										
	No	261 (91.90)	816 (93.90)	0.24									
EDI domain – not ready in communication skills	Yes	25 (8.80)	74 (8.52)										
	No	259 (91.20)	795 (91.48)	0.88									
Index of SES[a]	*N*	284	869		544	1297		475	1923		224	1060	
	Mean	0.25	0.03		0.16	0.10		0.16	0.09		0.38	0.32	
	Standard deviation	0.75	0.73	0.0001	0.76	0.85	0.11	0.70	0.78	<0.05	0.64	0.77	0.26
Mother's age at first birth (years)	*N*	284	869		544	1297		475	1923		224	1060	
	Mean	23.01	23.89		23.82	24.52		23.94	24.54		23.59	23.89	
	Standard deviation	5.39	5.45	<0.05	5.38	5.49	<0.05	5.48	5.47	<0.05	5.51	5.57	0.47
Child's age (months)	*N*	284	869		544	1297		475	1923		224	1060	
	Mean	62.66	62.75		62.88	62.70		62.86	62.66		62.71	62.75	
	Standard deviation	3.33	3.27	0.7	3.43	3.30	0.29	3.48	3.30	0.25	3.50	3.29	0.88

[a]Lower scores indicate higher SES.

matching). As expected, given the targeted nature of the intervention in school division 1, prior to matching, FDK students had significantly more challenges that could potentially affect their school performance than HDK students, including more students with: 2+ moves ($p < 0.0001$ for all outcomes), involvement in child welfare services ($p < 0.0001$ for grades 3–8 outcomes, $p < 0.05$ for grade 9 achievement), receipt of income assistance ($p < 0.0001$ for all outcomes), lower area-level income ($p < 0.0001$ for all outcomes), unmarried mothers ($p < 0.0001$ for all outcomes), younger age at first birth for mothers ($p < 0.0001$ for all outcomes), and higher SES index (indicating lower SES) ($p < 0.0001$) (Table A2).

Table 4 shows the standardised differences between the FDK and HDK cohorts in school divisions 1 and 2 for each of the outcomes, before and after the propensity score matching. Differences after matching are shown for the first match and for the multiple matches used in the models. Standardised differences of less than 20% (<0.20) were used to assess balance (Austin, 2009). As shown in the table, prior to the matching, the cohorts differed on most of the variables for all grade levels, with the exception of student age and sex, where differences were negligible. After matching, the cohorts showed no differences on the matching variables for grades 7, 8, and 9. For grade 3 outcomes, standardised differences were greater than 20% for income assistance (0.26) and the SES index (0.41) after matching. To account for differences that remained after matching, we used doubly robust estimation (Lunceford & Davidian, 2004). That is, first we matched children based on their estimated propensity score and then we modelled the relationship between FDK and school outcomes adjusting for area-level income to account for remaining differences in income assistance and SES. Doubly robust estimation has been shown to yield results robust to residual confounding (Bang, 2005; Lunceford & Davidian, 2004).

Table 5 shows the demographic information for students in school division 3. Students in FDK were significantly more likely to be from lower compared to higher income quintiles than students in HDK, and their mothers were less likely to be married. The FDK students also had mothers who were older at their first birth, and had a lower mean score on the SES index, indicating higher SES.

Table 6 (top panel) shows the regression results for the analysis of school divisions 1 and 2. Odds ratios (ORs) and 95% confidence intervals (CIs) for a priori contrasts for both kindergarten type by income quintile and for kindergarten type by sex interactions are shown. For grade 3 numeracy, among children living in the middle-income quintile (quintile 3), FDK children had three times the odds of meeting expectations for grade compared to HDK children. For grade 7 math, among students from the lowest income quintile, FDK students had two-and-a-half times the odds of meeting expectations for grade compared to HDK students, and FDK girls outperformed HDK girls. The results suggest that FDK benefits girls from the lowest income quintile in grade 7 math. For grade 7 engagement, FDK students from the highest income quintiles (quintiles 4 and 5 combined) were less likely to be engaged in their learning than HDK students in the same income quintiles. For grade 8 reading and writing, male FDK students were less likely to be meeting expectations than HDK males. No other differences were statistically significant. To summarise, out of 35 comparisons, five were statistically significant.

Table 6 (bottom panel) shows the regression results for the analysis of school division 3. The only statistically significant result was found for grade 7 engagement. FDK girls had three-and-a-half times the odds of being engaged in their learning as HDK girls. Out of 24 comparisons, only one was statistically significant.

Table 4. Standardised differences between school divisions 1 and 2 before and after propensity score match.

Outcomes	Grade	before PS matching	PS matching	PS matching (using 1st match only)
Contact with CFS				
	3	0.29	0.12	0.08
	7	0.24	0.13	0.00
	8	0.24	0.08	0.06
	9	0.13	0.02	0.01
Income assistance				
	3	0.56	0.24	0.04
	7	0.39	0.15	0.02
	8	0.43	0.12	0.07
	9	0.52	0.01	0.04
Mom's age at first birth				
	3	0.50	0.16	0.01
	7	0.42	0.13	0.01
	8	0.40	0.11	0.06
	9	0.41	0.00	0.03
Mom's marital status				
	3	0.42	0.10	0.02
	7	0.22	0.08	0.03
	8	0.26	0.03	0.04
	9	0.39	0.00	0.10
Number of moves				
	3	0.25	0.09	0.00
	7	0.22	0.07	0.01
	8	0.25	0.00	0.06
	9	0.37	0.00	0.02
SEFI 2				
	3	0.68	0.30	0.02
	7	0.54	0.17	0.03
	8	0.58	0.10	0.04
	9	0.87	0.10	0.03
Sex				
	3	0.09	0.07	0.04
	7	0.13	0.01	0.01
	8	0.09	0.05	0.00
	9	0.01	0.06	0.03
Age (months)				
	3	0.02	0.03	0.03
	7	0.07	0.03	0.01
	8	0.07	0.06	0.02
	9	0.02	0.01	0.06
EDI domain				
Language	3	0.06	0.10	0.08
Communication	3	0.00	0.03	0.02

Table 7 summarises the significant FDK effects for all outcomes and school divisions. Given that children were not randomly assigned to FDK *vs.* HDK, observed relationships between FDK and education outcomes may not reflect causal relationships; rather, they may be artefacts of residual confounding. We conducted sensitivity analyses to measure how large residual confounding would have to be to invalidate our

Table 5. Demographic characteristics of children in full-day (FDK) and half-day kindergarten (HDK), by outcomes examined, for school division 3.

		Grade 7			Grade 8			Grade 9		
Variables		FDK count (%)	HDK count (%)	*p*-value	FDK count (%)	HDK count (%)	*p*-value	FDK count (%)	HDK count (%)	*p*-value
Number of moves (years)	≤2	309 (92.24)	200 (90.09)		660 (93.35)	401 (90.72)		469 (92.14)	486 (92.40)	
	>2	26 (7.76)	22 (9.91)	0.38	47 (6.65)	41 (9.28)	0.10	40 (7.86)	40 (7.60)	0.88
Sex	Male	162 (48.36)	110 (49.55)		346 (48.94)	209 (47.29)		237 (46.56)	252 (47.91)	
	Female	173 (51.64)	112 (50.45)	0.78	361 (51.06)	233 (52.71)	0.59	272 (53.44)	274 (52.09)	0.66
Mother's marital status (at kindergarten start date)	Married	237 (70.75)	158 (71.17)		485 (68.60)	330 (74.66)		352 (69.16)	391 (74.33)	
	Not married	98 (29.25)	64 (28.83)	0.91	222 (31.40)	112 (25.34)	0.03	157 (30.84)	135 (25.67)	0.06
Income quintile	Q1 (lowest)	3 (0.90)	6 (2.70)		39 (5.52)	20 (4.52)		47 (9.23)	40 (7.60)	
	Q2	17 (5.07)	9 (4.05)		73 (7.69)	34 (7.69)		77 (15.13)	71 (13.50)	
	Q3	68 (20.30)	16 (7.21)		133 (18.81)	40 (9.05)		110 (21.61)	100 (19.01)	
	Q4	136 (40.60)	83 (37.39)		259 (36.63)	167 (37.78)		170 (33.40)	189 (35.93)	
	Q5 (highest)	111 (33.13)	108 (48.65)	<0.0001	203 (28.71)	181 (40.95)	<0.0001	105 (20.63)	126 (23.95)	0.42
Income assistance	Yes	28 (8.36)	14 (6.31)		58 (8.20)	33 (7.47)		35 (6.88)	48 (9.13)	
	No	307 (91.64)	208 (93.69)	0.37	649 (91.80)	409 (92.53)	0.65	474 (93.12)	478 (90.87)	0.18

Children in care/ protective services	Yes	25 (7.46)	17 (7.66)		45 (6.36)	34 (7.69)		26 (5.11)	29 (5.51)	
	No	310 (92.54)	205 (92.34)	0.93	662 (93.64)	408 (92.31)	0.39	483 (94.89)	497 (94.49)	0.77
Index of SES[a]	*N*	335	222		707	442		509	526	
	Mean	−0.27	−0.15		−0.35	−0.25		−0.28	−0.29	
	Standard deviation	0.51	0.42	<0.01	0.83	0.59	0.01	0.80	0.75	0.94
Mother's age at first birth (years)	*N*	335	220		707	440		508	523	
	Mean	24.65	25.40		25.85	25.55		26.26	25.81	
	Standard deviation	4.22	4.38	0.05	4.50	4.32	0.26	4.45	4.40	0.10
Child's age (months)	*N*	335	222		707	442		509	526	
	Mean	62.90	62.86		62.69	62.85		62.77	62.75	
	Standard deviation	3.45	3.42	0.89	3.30	3.30	0.43	3.22	3.27	0.92

[a]Lower scores indicate higher SES.

Table 6. Regression results for school divisions 1–3, comparing FDK to HDK students.

	Odds ratios (95% CI)					Estimate of mean difference
	Grade 3 reading	Grade 3 numeracy	Grade 7 math	Grade 7 engagement	Grade 8 read/write	Grade 9 index
Regression results for school divisions 1 and 2						
Income quintile 1, FDK *vs.* HDK	1.01 (0.52, 1.98)	1.09 (0.53, 2.23)	**2.47 (1.22, 5.02)**	1.05 (0.57, 1.94)	0.84 (0.46, 1.56)	0.14 (−0.22, 0.49)
Income quintile 2, FDK *vs.* HDK	1.62 (0.84, 3.14)	1.53 (0.79, 2.98)	1.63 (0.89, 2.99)	0.97 (0.59,1.59)	0.80 (0.51, 1.26)	−0.21 (−0.49, 0.07)
Income quintile 3, FDK *vs.* HDK	1.09 (0.55, 2.19)	**2.53 (1.29, 4.95)**	1.88 (0.99, 3.61)	1.38 (0.80, 2.38)	0.77 (0.44, 1.32)	−0.26 (−0.65, 0.12)[a]
Income quintiles 4 and 5, FDK *vs.* HDK	0.89 (0.41, 1.93)	1.14 (0.51, 2.54)	1.42 (0.74, 2.73)	**0.53 (0.28, 0.99)**	0.58 (0.32, 1.07)	
Males FDK *vs.* HDK	0.88 (0.51, 1.52)	1.45 (0.85, 2.49)	1.47 (0.87, 2.50)	0.86 (0.55, 1.34)	**0.64 (0.42, 0.997)**	−0.08 (−0.36, 0.19)
Females FDK *vs.* HDK	1.45 (0.81, 2.58)	1.50 (0.84, 2.70)	**2.23 (1.29, 3.87)**	1.01 (0.66, 1.55)	0.85 (0.57, 1.27)	−0.14 (−0.43, 0.15)
Regression results for school division 3						
Income quintile 1, FDK *vs.* HDK			0.73 (0.15, 3.61)[b]	3.99 (0.85, 18.75)[b]	1.52 (0.31, 7.30)	0.24 (−0.19, 0.67)
Income quintile 2, FDK *vs.* HDK					0.58 (0.18, 1.90)	−0.04 (−0.42, 0.34)
Income quintile 3, FDK *vs.* HDK					1.67 (0.58, 4.82)	−0.14 (−0.49, 0.21)
Income quintile 4, FDK *vs.* HDK			0.41 (0.10, 1.74)	1.61 (0.44, 5.85)	0.73 (0.34, 1.58)	0.10 (−0.18, 0.38)
Income quintile 5, FDK *vs.* HDK			0.83 (0.20, 3.42)	1.56 (0.41, 5.92)	0.80 (0.36, 1.74)	−0.08 (−0.41, 0.26)
Males FDK *vs.* HDK			0.55 (0.15, 2.06)	1.32 (0.39, 4.51)	0.73 (0.33, 1.65)	−0.06 (−0.33, 0.22)
Females FDK *vs.* HDK			0.72 (0.19, 2.69)	**3.53 (1.10, 11.32)**	1.28 (0.59, 2.77)	0.02 (−0.25, 0.28)

Bolded values indicate statistically significant differences between FDK and HDK.

[a]Income quintiles 3–5 combined.

[b]Income quintiles 1–3 combined.

Table 7. Summary of statistically significant FDK effects, and percent residual confounding effect required to invalidate results.

Assessment	Statistically significant effect	OR (95% CI)	Percent effect
Grade 3 numeracy, school divisions 1 and 2	Income quintile 3, FDK *vs.* HDK	2.53 (1.29, 4.95)	25.81%
Grade 7 math, school divisions 1 and 2	Income quintile 1, FDK *vs.* HDK	2.47 (1.22, 5.02)	29.17%
Grade 7 math, school divisions 1 and 2	Females, FDK *vs.* HDK	2.23 (1.29, 3.87)	31.61%
Grade 7 engagement, school divisions 1 and 2	Income quintiles 4 & 5, FDK *vs.* HDK	0.53 (0.28, 0.99)	1.39%
Grade 7 engagement, school division 3	Females, FDK *vs.* HDK	3.53 (1.10, 11.32)	7.62%
Grade 8 reading and writing, school divisions 1 and 2	Males FDK *vs.* HDK	0.64 (0.42, 0.997)	0.61%

results (Jiang & Foster, 2013; Rosenbaum, 2010) (shown by percent effect in last column of Table 7). For example, the percent effect for grade 7 math for school divisions 1 and 2 shows that the residual confounding would have to have an absolute effect on grade 7 math that is 31.6% of the observed relationship between FDK and grade 7 math; that is, there would need to be substantial residual confounding in order to nullify our findings. It is unlikely that after matching on the propensity score and using doubly robust estimation that such significant confounding remained. On the other hand, the percent effect for grade 8 reading and writing results for school divisions 1 and 2 shows that these results would be nullified by very tiny confounding – less than 1% of the relationship between FDK and reading. Thus, it is probable that the significant relationship between FDK and grade 8 reading may be attributable to residual confounding. The results for grade 7 engagement for the kindergarten type by income quintile interaction for school divisions 1 and 2 (1.4%) and the kindergarten type by sex interaction in school division 3 (7.6%) may also likely be attributable to residual confounding, whereas the sex by kindergarten type interaction for grade 7 math for school divisions 1 and 2 (31.6%) discussed above, as well as the income quintile by kindergarten type interaction for grade 7 math (29.2%) and the grade 3 numeracy in school divisions 1 and 2 (25.8%) are likely robust to residual confounding.

Discussion

This is the first longitudinal analysis examining long-term outcomes of FDK compared to HDK programmes beyond primary school in Canada. Using population-based administrative data we were able to follow children from kindergarten up to grade 9 and examine their performance on assessments in grades 3, 7, and 8 and their score on a grade 9 achievement index. Our study designs – propensity score matching and stepped-wedge – allow for stronger causal inferences than multivariate regression modeling (Brown & Lilford, 2006; Rubin, 2001), which has been used in many of the evaluations of FDK to date. Where FDK was introduced to all schools, we found only one statistically significant finding and determined that it was sensitive to unmeasured confounding; that is, the observed significant finding could easily become non-significant if

the model adjusted for all confounders. This finding confirms a relatively large literature from the USA that shows few if any longer-term academic benefits of FDK compared to HDK programmes (Cannon et al., 2006; Cooper et al., 2010; Olsen & Zigler, 1989; Votruba-Drzal et al., 2008; Wolgemuth et al., 2006). Where FDK was targeted to schools in low SES areas, out of six outcomes examined, we found three long-term FDK effects that were both statistically significant and robust to unmeasured confounding, although these significant effects were limited to specific sub-groups of students: improved grade 3 numeracy for middle-income students, improved grade 7 math for low-income students, and improved grade 7 math among girls. These findings of statistically significant long-term FDK effects are new and notable, but should be taken in the context of null findings for other sub-groups of students and for the other three outcomes examined for this school division.

The robust statistically significant findings for the targeted FDK programme were limited to assessments involving numeracy and math, rather than reading and writing, or school engagement. FDK students in this targeted FDK program who were in the middle-income group outperformed their HDK peers in grade 3 numeracy. And FDK students in this targeted FDK programme who were in the lowest income group and who were female outperformed their HDK peers in grade 7 math. Our analyses were not sufficiently powered to examine a three-way interaction, but the results suggest that targeted FDK programmes may engender some long-term benefits in mathematics for lower SES students, particularly girls.

Our finding of improved grade 7 performance in math for low-income girls is unique and requires further study. It will be important to follow these students further to determine whether scores in grade 12 standards exams in math are also elevated for FDK girls compared to HDK girls. If this finding is replicated in higher grades and in other jurisdictions, it will be critical to identify which aspects or components of FDK enhance lower income girls' numeracy skills and determine whether programming can then be tailored to enhance other types of skills (e.g. literacy) and to also improve boys' outcomes.

The fact that the few significant and robust results in our study were found only in the school division where FDK was targeted to low-income areas raises an important debate found in the literature; that is, whether early childhood programmes should be targeted or universal. Targeting programmes at children who are at the greatest risk of poor outcomes may seem like the most cost-effective way of improving population outcomes, since fewer children receive the (often costly) intervention and there may be a higher return on investment for lower income children (Cunha, Heckman, Lochner, & Masterov, 2006). However, targeting, by characteristics such as SES, ends up missing many in the population who are actually at risk of poor performance (Boivan & Hertzman, 2012) and may do little to improve overall population-level measures (McKinlay, 1998).

There may, of course, be other goals besides short- and long-term academic outcomes of FDK programmes. Heckman et al. (2013) have analysed data from the well-known Perry Preschool Program and demonstrated that it is not improvements in academic achievement per se that have led to the quite remarkable long-term positive effects of the programme, but changes in personality skills (including academic motivation). Thus, although the initial improvement in standardised test scores for children in the Perry Preschool Program faded over time, the programme was associated with other life-changing improvements including higher rates of completion of high school and lower involvement with the criminal justice system. Heckman et al.'s (2013) findings are consistent with a recent review from the Society for Research on Child

Development of 84 preschool programmes, which found that long-term benefits included more high school completion, fewer teen pregnancies and reduced crime rates for program participants, despite fading academic differences over the course of school between participants and non-participants (Yoshikawa et al., 2013).

Although it is tempting to use evidence about the long-term benefits of pre-kindergarten to estimate FDK effects (Hahn et al., 2014), caution should be exercised in generalising preschool programme results to FDK programmes. FDK programmes are not generally considered 'preschool' because they usually involve five-year-old children who will have access to some public schooling, whether half or full days in kindergarten. Furthermore, preschool programmes evincing long-term benefits often involve a parent training component (Yoshikawa et al., 2013), which FDK generally does not. Indeed, when the benefits of FDK programmes are compared head-to-head with preschool programmes, FDK programmes fare poorly (Reynolds & Temple, 2008). Reynolds and Temple (2008) reported that whereas programmes aimed at younger children (entry age into the programme one to four months through three years) yielded estimated benefit to cost ratios of 2:1 to over 10:1, FDK showed no benefit up to third grade. FDK is an expensive programme, particularly when implemented on a universal basis, and long-term outcomes will have to be considered carefully to determine whether the benefits outweigh the costs, or whether public funds would yield larger long-term payoffs if invested in programmes implemented earlier, during the preschool period.

To date there have been no long-term evaluations of FDK that have looked at outcomes beyond high school, focusing on the 'important societal outcomes' reported by Yoshikawa et al. (2013). It will be critical to extend the results of the current study past high school age to determine whether the children in the FDK programmes in Manitoba experience improvements in these societal outcomes, that is, whether they are more likely to complete high school, and whether FDK participation is associated with reductions in crime and teen pregnancy. With the capability of linking the population-based education data to both health and justice data in the MCHP Repository, we will have the opportunity to conduct this follow-up research within the next few years.

One of the driving factors in the proliferation of FDK programmes is the employment considerations of parents. Having kindergarten children in school for the full day rather than half a day can simplify child-care arrangements for working parents. Indeed, studies examining the universal, low-cost (first $5 and then $7 per day) child-care system in Quebec have shown that female employment and earnings increased significantly after implementation of the system (Baker, Gruber, & Milligan, 2008; Lefebvre & Merrigan, 2008). Whether a similar argument could be made for FDK remains to be determined. Arguments have also been made around the potential reduction of stress for parents whose children are in FDK programmes, due to not having to locate half-day child care or a means of delivering their children to off-site arrangements. It should be noted, however, that FDK does not replace the need for child care, as parents' hours of work do not necessarily coincide with the length of the school day. Likewise implementation of FDK can result in lost revenue for child-care centres. To remain financially viable, child-care centres may need to reorganise their spaces to serve only younger children, with the end result actually making it more difficult to find part-time spaces for before and after school and during school holidays for FDK-aged children (Manitoba Childcare Association, 2011).

A major strength of our analysis is the use of population-based administrative data for looking at long-term outcomes. The Repository housed at MCHP permits individual-level linkage across multiple data sets, allowing us to use student characteristic

information from multiple sources including mother's age at first birth, marital status, and number of residential moves, all factors that potentially influence student outcomes (Cutuli et al., 2013; Jutte et al., 2010; Strohschien et al., 2009). By linking the education data with the population registry, all children who attended FDK programmes could be followed regardless of whether they remained in the school or division where they attended kindergarten. It was also possible to examine large numbers of children providing ample power for statistical analyses. Furthermore, the anonymised nature of the data in the MCHP Repository circumvent the need for individual consent; therefore, the findings are not subject to participation bias. There are, however, limitations to our data sources that merit attention. Caution needs to be raised regarding the few robust statistically significant results we did find, given the number of comparisons that were made, and thus, increasing the chance of Type I error. Because very little research has been done on long-term outcomes for FDK, we chose not to adjust for multiple comparisons (e.g., Bonferonni correction factor) so that we could identify potential benefits associated with FDK that could be examined in future research. Additionally, although we were able to examine student engagement in grade 7, we were unable to examine other non-academic outcomes such as social and emotional development. Cooper et al. (2010) report some evidence that FDK may have positive effects on child independence, self-confidence and ability to play with others, but negative effects on attitudes towards school and behaviour; however, caution is raised about the non-experimental nature of the studies in their review. Clearly, examining both short- and long-term social and emotional outcomes of FDK should be a priority for future research. Data were also unavailable on programme quality, student attendance, or activities that HDK students participated in for the remainder of their school day. All these factors could potentially affect the longer-term outcomes of students. Furthermore, the assessments used in grades 3, 7, and 8 were teacher administered and therefore subject to teacher biases regarding students. Whilst it is entirely possible that teachers conducting grade 3 assessments knew whether their students had attended FDK or HDK programmes, and thus their own biases about the benefits of these programmes may have entered into their evaluations, by grades 7 and 8 the students would be attending different schools than where they attended kindergarten, and therefore the likelihood that the teachers conducting the assessments were aware of their kindergarten history is unlikely. It is also unlikely that teachers realised when they were conducting their assessments, that these would later be used in an evaluation of the FDK programmes. It is important to note that drawing causal inferences about results produced in a propensity score matched cohort study rest on the strong assumption of ignorable assignment to FDK; that is, after matching on the estimated propensity score, we have controlled all observed and unobserved confounding. This strong assumption may not hold in our study; however, we were able to test how sensitive any statistically significant results were to violations of this assumption.

Conclusion

Our findings indicate no apparent benefits of universal FDK, but there may be some benefits from targeted FDK programmes, under certain circumstances. FDK programmes targeted at low-income areas showed long-term improvements in numeracy for low-income girls. The evidence from this study suggests that expectations for population-level improvements in long-term academic performance for FDK programmes may be overstated.

Acknowledgements

This study is part of a program of research being conducted by the PATHS Equity Team: James Bolton, Marni Brownell, Charles Burchill, Elaine Burland, Mariette Chartier, Dan Chateau, Malcolm Doupe, Greg Finlayson, Randall Fransoo, Chun Yan Goh, Milton Hu, Doug Jutte, Alan Katz, Laurence Katz, Lisa Lix, Patricia J. Martens, Colleen Metge, Nathan C. Nickel, Colette Raymond, Les Roos, Noralou Roos, Rob Santos, Joykrishna Sarkar, Mark Smith, Carole Taylor, Randy Walld. The authors would like to thank Theresa Daniuk for assistance with manuscript preparation. The authors acknowledge the Manitoba Centre for Health Policy for use of data contained in the Population Health Research Data Repository under project # 2012 - 006 (HIPC #2011/2012 – 24G). The results and conclusions are those of the authors and no official endorsement by MCHP, Manitoba Health, or other data providers is intended or should be inferred. Data used in this study are from the Population Health Research Data Repository housed at MCHP, University of Manitoba, and were derived from data provided by Healthy Child Manitoba; Manitoba Education and Advanced Learning; Manitoba Family Services; Manitoba Health; Manitoba Jobs and the Economy; and Statistics Canada.

Funding

This research was funded by the Canadian Institutes of Health Research (CIHR) [grant number ROH - 115206] and the Heart & Stroke Foundation of Canada [grant number PG-12-0534], under the programme of research entitled 'PATHS Equity for Children: a program of research into what works to reduce the gap for Manitoba's children.' Dr Marni Brownell acknowledges the financial support of the Government of Manitoba through the Manitoba Center for Health Policy Population-Based Child Health Research Award. Dr. Patricia Martens wishes to acknowledge funding from the Canadian Institutes of Health Research (CIHR) and the Public Health Agency of Canada (PHAC) for her CIHR/PHAC Applied Public Health Research Chair (2008–2014). Dr Alan Katz acknowledges the support of the Manitoba Health Research Council and the Heart and Stroke Foundation for his Research Chair in Primary Prevention (2013–2018).

Notes on contributors

M.D. Brownell is a Senior Research Scientist at the Manitoba Centre for Health Policy, and Associate Professor in the Department of Community Health Sciences, College of Medicine, Faculty of Health Sciences, University of Manitoba.

N.C. Nickel is a Research Scientist at the Manitoba Centre for Health Policy, and Assistant Professor in the Department of Community Health Sciences, College of Medicine, Faculty of Health Sciences, University of Manitoba.

D. Chateau is a Research Scientist at the Manitoba Centre for Health Policy, and Assistant Professor in the Department of Community Health Sciences, College of Medicine, Faculty of Health Sciences, University of Manitoba.

P.J. Martens is a Senior Research Scientist and former Director of the Manitoba Centre for Health Policy, Distinguished Professor in the Department of Community Health Sciences, College of Medicine, Faculty of Health Sciences, University of Manitoba, and Member of the Order of Canada.

C. Taylor is a Data Analyst with the Manitoba Centre for Health Policy.

L. Crockett is a Research Assistant with the Manitoba Centre for Health Policy, and an MSc student in the Department of Community Health Sciences, College of Medicine, Faculty of Health Sciences, University of Manitoba.

A. Katz is a Senior Research Scientist and Acting Director of the Manitoba Centre for Health Policy, and Professor in the Departments of Community Health Sciences and Family Medicine, College of Medicine, Faculty of Health Sciences, University of Manitoba.

J. Sarkar is a Data Analyst with the Manitoba Centre for Health Policy.

E. Burland is the Research Manager at the Manitoba Centre for Health Policy.

C.Y. Goh is a Research Assistant at the Manitoba Centre for Health Policy.

The PATHS Equity Team is a multi-disciplinary research team focused on understanding what works to reduce inequity in children's outcomes. http://umanitoba.ca/faculties/medicine/units/community_health_sciences/departmental_units/mchp/8383.html

References

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*, 3083–3107. doi:10.1002/sim.3697.

Baker, M., Gruber, J., & Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, *116*, 709–745.

Bang, H. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, *61*, 962–972.

Bingham, G. E., & Hall-Kenton, K. M. (2012). Full- and half-day kindergarten programmes: Examining impacts on second language learners. *Early Child Development and Care*, *183*(2), 185–199.

Boivan, M., & Hertzman, C. (Eds.) (2012). *Early Childhood Development: Adverse Experiences and Developmental Health*. Ottawa, ON: Royal Society of Canada.

Brown, C. A., & Lilford, R. J. (2006). The stepped wedge trial design: A systematic review. *BMC Medical Research Methodology*, *6*, 54, doi:10.1186/1471-2288-6-54.

Brownell, M., Mayer, T., Martens, P. J., Kozyrskyj, A., Fergusson, P., Bodnarchuk, J., … Derksen, S. (2002). Using a population-based health information system to study child health. *Canadian Journal of Public Health*, *93*(Suppl 2), S9–S14.

Brownell, M. D., Roos, N. P., Fransoo, R., Roos, L. L., Guèvremont, A., MacWilliam, L., … Yallop, L. (2006). Is the class half-empty? A population-based perspective on socio-economic status and educational outcomes. *IRPP Choices (Institute for Research on Public Policy)*, *12*(5), 1–30.

Brownell, M., Roos, N., MacWilliam, L., Leclair, L., Ekuma, O., & Fransoo, R. (2010). Academic and social outcomes for high-risk youths in Manitoba. *Canadian Journal of Education*, *33*, 804–836.

Cannon, J. S., Jacknowitz, A., & Painter, G. (2006). Is full better than half? Examining the longitudinal effects of full-day kindergarten attendance. *Journal of Policy Analysis and Management*, *2*, 299–321.

Chang, M. (2012). Academic performance of language-minority students and all-day kindergarten: A longitudinal study. *School Effectiveness and School Improvement*, *23*(1), 21–48.

Chang, M., & Singh, K. (2008). Is all-day kindergarten better for children's academic performance? Evidence from the early childhood longitudinal study. *The Australian Journal of Early Childhood*, *33*(4), 35–42.

Chase-Lansdale, P., Coley, R., Lohman, B., & Pittman, L. (2002). *Welfare reform: What about the children? Welfare, children & families: A three-city study. Policy brief*. ERIC Document Reproduction Service No. ED464169, John Hopkins University, Baltimore, MD. IPR Policy Briefs, 1(3), Institute for Policy Research, Northwestern University, May 2002. http://www.ipr.northwestern.edu/publications/docs/policybriefs/lansdalebrief.pdf

Chateau, D., Metge, C., Prior, H., & Soodeen, R. A. (2012). Learning from the census: The Socio-Economic Factor Index (SEFI) and health outcomes in Manitoba. *Canadian Journal of Public Health*, *103*(2 Suppl.), S223–S27.

Cooper, H., Allen, A., Patall, E. A., & Dent, A. L. (2010). Effects of full-day kindergarten on academic achievement and social development. *Review of Educational Research*, *80*, 34–70.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence of life cycle skill formation. In E. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education* (pp. 697–812). Amsterdam: North-Holland.

Cutuli, J. J., Desgardins, C. D., Herbers, J. E., Long, J. D., Heistad, D., Chan, C. K., … Hinz, E. (2013). Academic achievement trajectories of homeless and highly mobile students: Resilience in the context of chronic and acute risk. *Child Development*, *84*, 841–857.

D'Agostino, R. B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, *17*, 2265–2281.

DeCicca, P. (2007). Does full-day kindergarten matter? Evidence from the first two years of schooling. *Economics of Education Review*, *26*, 67–82.

Elicker, J., & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly*, *12*, 459–480.

Forget-Dubois, N., Lemelin, J.-P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., & Tremblay, R. E. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education and Development*, *18*, 405–426.

Geenen, S., & Powers, L. (2006). Are we ignoring youths with disabilities in foster care? An examination of their school performance. *Social Work*, *51*, 233–241.

Guèvremont, A., Roos, N. P., & Brownell, M. (2007). Predictors and consequences of grade retention: Examining data from Manitoba, Canada. *Canadian Journal of School Psychology*, *22*(1), 50–67.

Guo, S., & Fraser, M. W. (2009). *Propensity score analysis: Statistical methods and applications*. Thousand Oaks, CA: Sage Publications.

Hahn, R. A., Rammohan, V., Truman, B. I., Milstein, B., Johnson, R. L., Muntaner, C., … the Community Preventive Services Task Force. (2014). Effects of full-day kindergarten on the long-term health prospects of children in low-income and racial/ethnic-minority populations: A community guide systematic review. *American Journal of Preventive Medicine*, *46*, 312–323.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*, 2052–2086.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945–960.

Ingersoll, G. M., Scamman, J. P., & Eckerling, W. D. (1989). Geographic mobility and student achievement in an urban setting. *Educational Evaluation and Policy Analysis*, *11*, 143–149.

Jaffee, S. R. (2002). Pathways to adversity in young adulthood among early childbearers. *Journal of Family Psychology*, *16*(1), 38–49.

Janus, M., & Offord, D. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioral Science*, *39*(1), 1–22.

Jiang, M., & Foster, E. M. (2013). Duration of breastfeeding and childhood obesity: A generalized propensity score approach. *Health Services Research*, *48*, 628–651. doi:10.1111/j.1475-6773.2012.01456.x.

Jutte, D., Roos, N., Brownell, M., Briggs, G., MacWilliam, L., & Roos, L. L. (2010). The ripples of adolescent motherhood: Social, educational and medical outcomes for children of teen and prior teen moms. *Academic Pediatrics*, *10*, 293–301. doi:10.1016/j.acap.2010.06.008.

Lee, V. E., Burkam, D. T., Ready, D., Honigman, J., & Meisels, S. (2006). Full-day versus half-day kindergarten: In which program do children learn more? *American Journal of Education*, *112*, 163–208.

Lefebvre, P., & Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from Canada. *Journal of Labor Economics*, *26*, 519–548.

Lunceford, J. K., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine*, *23*, 2937–2960.

Manitoba Childcare Association. (2011). *Planning for the future: Nursery and full-day kindergarten in public schools*. Retrieved from http://www.mccahouse.org/index.htm

McKinlay, J. B. (1998). Paradigmatic obstacles to improving the health of population: Implications for health policy. *Salud Publica Mex*, *40*, 369–379.

Morgan, S. L., & Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York, NY: Cambridge University Press.

Olsen, D., & Zigler, E. (1989). An assessment of the all-day kindergarten movement. *Early Childhood Research Quarterly*, *4*(2), 167–186.

Oreopoulos, P., Stabile, M., Walld, R., & Roos, L. L. (2008). Short-, medium-, and long-term consequences of poor infant health: An analysis using siblings and twins. *Journal of Human Resources*, *43*(1), 88–138.
Pascal, C. E. (2009). *With our best future in mind: Early learning in Ontario*. Toronto, ON: Government of Ontario.
Preston, J. P., Cottrell, M., Pelletier, T. R., & Pearce, J. V. (2012). Aboriginal early childhood education in Canada: Issues of context. *Journal of Early Childhood Research*, *10*, 3–18.
Puleo, V. T. (1988). A review and critique of research on full-day kindergarten. *The Elementary School Journal*, *88*, 427–439.
Reynolds, A. J., & Temple, J. A. (2008). Cost-effective early childhood development programs from preschool to third grade. *Annual Review of Clinical Psychology*, *4*, 109–139. doi:10.1146/annurev.clinpsy.3.022806.091411.
Roos, N. P., Brownell, M., Guèvremont, A., Fransoo, R., Levin, B., MacWilliam, L., & Roos, L. L. (2006). The complete story: A population-based perspective on school performance and educational testing. *Canadian Journal of Education*, *29*(3), 1–22.
Roos, L. L., Gupta, S., Soodeen, R. A., & Jebamani, L. (2005). Data quality in an information-rich environment: Canada as an example. *Canadian Journal on Aging*, *24*(Suppl. 1), 153–170.
Roos, L. L., Hiebert, B., Manivong, P., Edgerton, J., Walld, R., MacWilliam, L., & de Rocquigny, J. (2013). What is most important: Social factors, health selection, and adolescent educational achievement. *Social Indicators Research*, *110*, 385–414.
Roos, L. L., & Nicol, J. P. (1999). A research registry: Uses, development, and accuracy. *Journal of Clinical Epidemiology*, *52*(1), 39–47.
Rosenbaum, P. R. (2010). *Observational Studies*. New York, NY: Springer + Business Media.
Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*, 169–188.
Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, *100*, 322–331.
Rubin, D. B. (2008). For objective causal inference, Design trumps analysis. *Annals of Applied Statistics*, *2*, 808–840.
Rubin, D. B., & Thomas, D. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, *52*, 249–264.
Scanlon, E., & Devine, K. (2001). Residential mobility and youth well-being: Research, policy, and practice issues. *Journal of Sociology and Social Welfare*, *28*, 119–138.
Scherr, T. (2007). Educational experiences of children in foster care: Metaanalyses of special education, retention and discipline rates. *School Psychology International*, *28*, 419–436.
Schroeder, J. (2007). Full-day kindergarten offsets negative effects of poverty on state tests. *European Early Childhood Education Research Journal*, *15*, 427–439.
Strohschein, L., Roos, N., & Brownell, M. (2009). Family structure histories and high school completion: Evidence from a population based registry. *Canaian Journal of Sociology*, *34*(1), 83–103.
Votruba-Drzal, E., Li-Grining, C. P., & Maldonado-Carreno, C. (2008). A developmental perspective on full- versus part-day kindergarten and children's academic trajectories through fifth grade. *Child Development*, *79*, 957–978. doi:10.1111/j.1467-8624.2008.01170.x.
Walston, J. T., & West, J. (2004). *Full-day and half-day kindergarten in the United States: Findings from the early childhood longitudinal study, kindergarten class of 1998–99 (NCES 2004–078)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, U.S. Government Printing Office.
Wolgemuth, J. R., Cobb, R. B., Winokur, M. A., Leech, N., & Ellerby, D. (2006). Comparing longitudinal academic achievement of full-day and half-day kindergarten students. *The Journal of Educational Research*, *99*, 260–270.
Yan, W., & Lin, Q. (2005). Effects of class size and length of day on kindergartners' academic achievement: Findings from early childhood longitudinal study. *Early Education & Development*, *16*, 49–68.
Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., … Ludwig, J. (2013). *Investing in our future: The evidence base on preschool education*. Ann Arbor, MI: Society for Research on Child Development. Retrieved October 2013 from http://www.srcd.org/sites/default/files/documents/washington/mb_2013_10_16_investing_in_children.pdf

Appendix

Table A1. Description of outcomes examined.

Outcome	Competencies assessed	Level achieved
Grade 3 reading	Reflects on and sets reading goals	Needs ongoing help
	Uses strategies during reading to make sense of texts	Approaching expectations
	Demonstrates comprehension	Meeting expectations
Grade 3 numeracy	Predicts an element in a repeating pattern	Needs ongoing help
	Understands that the equal symbol represents an equality of the terms found on either side of the symbol	Approaching expectations
	Understands that a given whole number may be represented in a variety of ways (to 100)	Meeting expectations
	Uses various mental math strategies to determine answers to addition and subtraction questions to 18	
Grade 7 math	Orders fractions	Emerging skills
	Orders decimal numbers	Developing skills
	Student understands that a given number may be represented in a variety of ways	Established skills
	Uses number patterns to solve mathematical problems	
	Uses a variety of strategies to calculate and explain a mental math problem	
Grade 7 student engagement	Demonstrates and interest in his/her learning	Emerging skills
	Engages in self-assessment	Developing skills
	Is aware of learning goals of a unit of study and/or personal learning goals	Established skills
	Participates in lessons	Inconsistent skills
	Accepts responsibility for assignments	
Grade 8 reading and writing	Understands key ideas and messages in a variety of texts	Not meeting expectations
	Interprets a variety of texts	Approaching expectations
	Responds critically to a variety of texts	Meeting expectations
	Generates selects and organises ideas to support the readers' understanding	
	Chooses language (word choices, sentence patterns) to make an impact on the reader	
	Uses conventions (spelling, grammar and/or punctuation) and resources to edit and proofread to make meaning clear	
Grade 9 index	The index was developed based on possible average marks in all classes and the number of credits earned during the grade 9 school year. A scaled logit score was created based on the rank categories each individual was assigned to (Roos et al., 2013)	

Table A2. Demographic characteristics of children in FDK and HDK, by outcomes for the entire sample prior to matching, school divisions 1 and 2.

			Grade 3				Grade 7				Grade 8				Grade 9			
			Before matching		PS matching		Before matching		PS matching		Before matching		PS matching		Before matching		PS matching	
Descriptive of kids			FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)	FDK count (%)	HDK count (%)
Number of moves (years)	≤2		210 (71.67)	1261 (82.04)	206 (72.54)	658 (75.72)	416 (74.42)	1749 (82.81)	406 (74.63)	1000 (77.10)	353 (72.78)	3592 (82.94)	347 (73.05)	1409 (73.27)	148 (65.78)	3510 (81.95)	148 (66.07)	722 (68.11)
	>2		83 (28.33)	276 (17.96)	78 (27.46)	211 (24.28)	143 (25.58)	363 (17.19)	138 (25.37)	297 (22.90)	132 (27.22)	739 (17.06)	128 (26.95)	514 (26.73)	77 (34.22)	773 (18.05)	76 (33.93)	338 (31.89)
		p-value		<0.0001		0.28		<0.0001		0.26		<0.0001		0.92		<0.0001		0.55
Sex	Male		164 (55.97)	788 (51.27)	158 (55.63)	468 (53.86)	314 (56.17)	1064 (50.38)	306 (56.25)	703 (54.20)	266 (54.85)	2174 (50.20)	259 (54.53)	1003 (52.16)	112 (49.78)	2157 (50.36)	112 (50.00)	554 (52.26)
	Female		129 (44.03)	749 (48.73)	126 (44.37)	401 (46.14)	245 (43.83)	1048 (49.62)	238 (43.75)	594 (45.80)	219 (45.15)	2157 (49.80)	216 (45.47)	920 (47.84)	113 (50.22)	2126 (49.64)	112 (50.00)	506 (47.74)
		p-value		0.14		0.60		<0.05		0.42		0.05		0.35		0.86		0.54
Mother's marital status (at kindergarten start date)	Married		98 (33.68)	827 (53.98)	98 (34.51)	342 (39.36)	274 (49.28)	1271 (60.35)	273 (50.18)	692 (53.35)	236 (48.86)	2672 (61.88)	235 (49.47)	982 (51.07)	99 (44.00)	2693 (63.16)	99 (44.20)	483 (45.57)
	Not married		193 (66.32)	705 (46.02)	186 (65.49)	527 (60.64)	282 (50.72)	835 (39.65)	271 (49.82)	605 (46.65)	247 (51.14)	1646 (38.12)	240 (50.53)	941 (48.93)	126 (56.00)	1571 (36.84)	125 (55.80)	577 (54.43)
		p-value		<0.0001		0.14		<0.0001		0.21		<0.0001		0.53		<0.0001		0.71
Income quintile	Q1 (lowest)		87 (29.69)	161 (10.47)	82 (28.87)	140 (16.11)	127 (22.72)	287 (13.59)	120 (22.06)	243 (18.74)	109 (22.47)	421 (9.72)	102 (21.47)	331 (17.21)	65 (28.89)	417 (9.74)	64 (28.57)	249 (23.49)
	Q2		89 (30.38)	310 (20.17)	88 (30.99)	249 (28.65)	205 (36.67)	499 (23.63)	198 (36.40)	363 (27.99)	192 (39.59)	903 (20.85)	190 (40.00)	550 (28.60)	116 (51.56)	914 (21.34)	116 (51.79)	367 (34.62)
	Q3		70 (23.89)	410 (26.68)	68 (23.94)	239 (27.50)	128 (22.90)	409 (19.37)	127 (23.35)	254 (19.58)	106 (21.86)	1033 (23.85)	105 (22.11)	463 (24.08)	31 (13.78)	1037 (24.21)	31 (13.84)	252 (23.77)
	Q4		33 (11.26)	503 (32.73)	33 (11.62)	203 (23.36)	67 (11.99)	576 (27.27)	67 (12.32)	299 (23.05)	53 (10.93)	1294 (29.88)	53 (11.16)	429 (22.31)	9 (4.00)	1261 (29.44)	9 (4.02)	150 (14.15)
	Q5 (highest)		14 (4.78)	153 (9.95)	13 (4.58)	38 (4.37)	32 (5.72)	341 (16.15)	32 (5.88)	138 (10.64)	25 (5.15)	680 (15.70)	25 (5.26)	150 (7.80)	4 (1.78)	654 (15.27)	4 (1.79)	42 (3.96)
		p-value		<0.0001		<0.0001		<0.0001		<0.0001		<0.0001		<0.0001		<0.0001		<0.0001
Income assistance	Yes		124 (42.32)	269 (17.50)	119 (41.90)	259 (29.80)	208 (37.21)	438 (20.74)	196 (36.03)	381 (29.38)	189 (38.97)	851 (19.65)	182 (38.32)	627 (32.61)	99 (44.00)	881 (20.57)	98 (43.75)	437 (41.23)
	No		169 (57.68)	1268 (82.50)	165 (58.10)	610 (70.20)	351 (62.79)	1674 (79.26)	348 (63.97)	916 (70.62)	296 (61.03)	3480 (80.35)	293 (61.68)	1296 (67.39)	126 (56.00)	3402 (79.43)	126 (56.25)	623 (58.77)
		p-value		<0.0001		<0.001		<0.0001		<0.01		<0.0001		<0.05		<0.0001		0.49

Children in care/ protective services	Yes		62 (21.16)	166 (10.80)	59 (20.77)	139 (16.00)	124 (22.18)	304 (14.39)	116 (21.32)	227 (17.50)	105 (21.65)	553 (12.77)	98 (20.63)	340 (17.68)	41 (18.22)	582 (13.59)	41 (18.30)	194 (18.30)
	No		231 (78.84)	1371 (89.20)	225 (79.23)	730 (84.00)	435 (77.82)	1808 (85.61)	428 (78.68)	1070 (82.50)	380 (78.35)	3778 (87.23)	377 (79.37)	1583 (82.32)	184 (81.78)	3701 (86.41)	183 (81.70)	866 (81.70)
		p-value		<0.0001		0.06		<0.0001		0.05		<0.0001		0.14		<0.05		1.00
EDI domain – not ready in physical health	Yes		56 (19.11)	136 (8.85)	53 (18.66)	90 (10.36)												
	No		237 (80.89)	1401 (91.15)	231 (81.34)	779 (89.64)												
		p-value		<0.0001		<0.001												
EDI domain – not ready in emotional maturity	Yes		47 (16.04)	178 (11.58)	45 (15.85)	96 (11.05)												
	No		246 (83.96)	1359 (88.42)	239 (84.15)	773 (88.95)												
		p-value		<0.05		<0.05												
EDI domain – not ready in social competence	Yes		38 (12.97)	133 (8.65)	36 (12.68)	72 (8.29)												
	No		255 (87.03)	1404 (91.35)	248 (87.32)	797 (91.71)												
		p-value		<0.05		<0.05												
EDI domain – not ready in language development	Yes		25 (8.53)	108 (7.03)	23 (8.10)	53 (6.10)												
	No		268 (91.47)	1429 (92.97)	261 (91.90)	816 (93.90)												
		p-value		0.36		0.24												
EDI domain – not ready in communication skills	Yes		27 (9.22)	143 (9.30)	25 (8.80)	74 (8.52)												
	No		266 (90.78)	1394 (90.70)	259 (91.20)	795 (91.48)												
		p-value		0.96		0.88												

(*Continued*)

Table A2. Continued.

			Grade 3				Grade 7				Grade 8				Grade 9			
			Before matching		PS matching		Before matching		PS matching		Before matching		PS matching		Before matching		PS matching	
			FDK	No FDK	FDK	No FDK	FDK	No FDK	FDK	No FDK	FDK	No FDK	FDK	No FDK	FDK	No FDK	FDK	No FDK
Socioeconomic factor index (SEFI) 2	*N*		292	1534	284	869	559	2112	544	1297	485	4330	475	1923	225	4283	224	1060
	Mean		0.26	−0.25	0.25	0.03	0.18	−0.11	0.16	0.10	0.18	−0.25	0.16	0.09	0.39	−0.25	0.38	0.32
	Standard deviation		0.75	0.74	0.75	0.73	0.77	0.85	0.76	0.85	0.72	0.79	0.70	0.78	0.65	0.80	0.64	0.77
	Minimum		−2.34	−2.27	−2.34	−2.04	−3.73	−2.51	−3.73	−2.51	−1.49	−3.94	−1.49	−2.51	−1.49	−3.94	−1.49	−1.73
	Maximum		2.99	4.57	2.99	4.57	3.41	3.55	3.41	3.55	3.41	3.55	3.41	3.55	2.38	3.55	2.38	3.41
		p-value		<0.0001		<0.0001		<0.0001		0.11		<0.0001		<0.05		<0.0001		0.26
Mother's age at first birth (years)	*N*		291	1533	284	869	556	2110	544	1297	483	4322	475	1923	225	4268	224	1060
	Mean		22.92	25.67	23.01	23.89	23.70	25.79	23.82	24.52	23.84	26.00	23.94	24.54	23.58	25.78	23.59	23.89
	Standard deviation		5.37	5.51	5.39	5.45	5.39	5.59	5.38	5.49	5.50	5.31	5.48	5.47	5.50	5.26	5.51	5.57
	Minimum		15.00	14.00	15.00	15.00	14.00	13.00	14.00	13.00	14.00	14.00	14.00	14.00	14.00	14.00	14.00	14.00
	Maximum		42.00	43.00	42.00	43.00	39.00	45.00	39.00	45.00	39.00	45.00	39.00	45.00	39.00	45.00	39.00	45.00
		p-value		<0.0001		<0.05		<0.0001		<0.05		<0.0001		<0.05		<0.0001		0.47
Child's age (months)	*N*		293	1537	284	869	559	2112	544	1297	485	4331	475	1923	225	4283	224	1060
	Mean		62.67	62.61	62.66	62.75	62.85	62.59	62.88	62.70	62.87	62.63	62.86	62.66	62.68	62.63	62.71	62.75
	Standard deviation		3.33	3.25	3.33	3.27	3.44	3.33	3.43	3.30	3.47	3.31	3.48	3.30	3.50	3.34	3.50	3.29
	Minimum		57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00	57.00
	Maximum		68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00
		p-value		0.77		0.70		0.11		0.29		0.15		0.25		0.83		0.88