Higher Education Quality Council of Ontario

An agency of the Government of Ontario

The Power of Connected Data: Charting Student Pathways to and through Postsecondary in Hamilton

April Au, Jackie Pichette and Karen Robson

Published by:

The Higher Education Quality Council of Ontario

88 Queens Quay West, Suite 2500
Toronto, ON
Canada, M5J 0B8

Phone: (416) 212-3893
Fax: (416) 212-3899
Web: www.heqco.ca

E-mail: info@heqco.ca

Cite this publication in the following format:

Au, A., Pichette, J., & Robson, K. (2022) *The Power of Connected Data : Charting Student Pathways to and through Postsecondary in Hamilton*. Toronto: Higher Education Quality Council of Ontario

Higher Education Quality Council of Ontario

An agency of the Government of Ontario

Acknowledgements

This report would not have been possible without the Hamilton CRP Leadership Committee members who guided the development and use of the CRP dataset. Nor would it have been possible without the inspired thinking and contributions of Fiona Deller, Lauren Hudak, Lorraine Valmadrid, Sarah Glen and Matt Goodman.

We are also grateful to the many HEQCO staff members who supported the development of the CRP dataset and contributed research to this report.

Table of Contents

List of Tables

List of Figures

Executive Summary

The Hamilton Community Research Partnership (CRP) is a data-sharing coalition that includes six partners: the Hamilton-Wentworth public and Catholic school boards, McMaster University, Mohawk College, the Hamilton Community Foundation and the Higher Education Quality Council of Ontario (HEQCO).

The CRP members joined together with three goals. First, they sought to better understand the education pathways of Hamilton students — and specifically, pathways from secondary school to and through college and university.[1] Before the CRP, most available research about barriers to postsecondary access in Ontario came from the Toronto District School Board (TDSB): a school board with a uniquely large, diverse and urban jurisdiction, whose research findings may not necessarily hold true for other communities.

The CRP's second goal was to develop and test a data-sharing mechanism using the Ontario Education Number (OEN). The OEN offers the potential to connect datasets that track students' educational pathways over time while minimizing the risk of identifying those students. Currently, access to OEN-linked data is restricted to the Ontario government (Gallagher-Mackay, 2017), but institutions with data tied to the OEN are free to link their data with willing partners and available data sources. For example, the TDSB has used the OEN numerous times to link their own data with other sources, namely: with a single Ontario university; with college and university application centres; and with Statistics Canada's Education and Labour Market Linkage Platform. To our knowledge, no Ontario community has used the OEN to link administrative data across multiple school boards, postsecondary partners and application centres to produce one dataset. If the CRP data-sharing and matching process proved to be effective, the CRP could offer a proof of concept to other Ontario communities to examine educational pathways for themselves.

Finally, the CRP members came together with a goal of building trust and setting up procedures for ongoing, collaborative educational research in their community. The partners refer to the process of assembling the dataset and using it to produce this report as "the pilot phase" of a long-term partnership. Data security and privacy protection were key to building trust during this phase. To ensure privacy, the partners designated only one Partnership Data Analyst (PDA) to have direct access to the dataset once it was

[1] Other pathways acknowledged by the Ministry of Education, including apprenticeship training, entering the workplace, and community-living, were outside of the scope of this project.

assembled, and this sole individual shared aggregate results with the authors of this report using encrypted emails.

Our pilot accomplished each of the partners' goals. The CRP successfully linked de-identified datasets from both school boards, McMaster University, Mohawk College and Ontario's application centres using the OEN and probabilistic matching. The resulting dataset has a 99.6% match rate, meaning we were able to connect school board and postsecondary data successfully for nearly all students who entered Grade 9 in 2010 in Hamilton and who continued their postsecondary education at either of the local, publicly assisted postsecondary institutions — the first dataset of its kind in Ontario.

Once the dataset was assembled, we conducted statistical analyses to answer the following questions: Who graduates high school in Hamilton? Who goes to college or university in Ontario? And for those who study in Hamilton, who completes college or university programs? In total, we examined pathways for 5,310 students, contributing to a growing body of research into the factors affecting educational outcomes such as high school graduation and postsecondary access.

We found that the same factors shown to predict postsecondary access most reliably in other parts of North America apply to Hamilton and may have broad applicability across Ontario — something researchers and educators have been reluctant to assume. Accounting for other sociodemographic factors, our analysis confirmed Grade 9 credit accumulation and secondary grades are the strongest determinants of graduation at both secondary and postsecondary levels. They were also closely tied to other demographic factors like neighbourhood income, secondary stream and exceptionality.

Two demographic variables provided particularly striking advantages along Hamilton educational pathways: income and secondary stream. Students from higher-income neighbourhoods were more likely to graduate high school, confirm a university offer and pursue STEM programs. Students who pursued the academic stream in high school gained similar advantages. Aligned with findings from the TDSB, we found students in non-academic courses were less likely to graduate high school or attend postsecondary than their peers in academic courses. Further, we found students from the applied stream were less likely to confirm a postsecondary offer of admission, and to university in particular, relative to students in the academic stream. We also found students who were in the academic stream in high school graduated both college and university in higher numbers.

Despite yielding these important findings, the CRP dataset falls short of fully answering the questions posed in this report. This pilot phase of the project relied on historic data (for students who entered Grade 9 for the first time during the 2010-11 academic year), collected prior to the Ontario government's requirement that school boards collect race-

based data (Ontario Newsroom, 2020). Privacy was also paramount, which meant some available variables needed to be suppressed or excluded to eliminate the risk of identifying students. As such, our dataset does not currently allow for analysis of sociodemographic factors like race, sexuality or Indigenous identity in relation to outcomes like graduation. This severely limits our ability to inform equity and inclusion strategies. By providing 'notice of collection,' the partners have positioned themselves to improve future iterations of the dataset. But even with the inclusion of nuanced variables, a community-level dataset will fall short; a provincial-level data infrastructure would allow us to answer the research questions addressed in this report with more completeness and certainty. With this in mind, we offer the following recommendations to the provincial government:

- Require that institutions collect consistent administrative and demographic data[2] tied to the OEN.
- Use the OEN to build a longitudinal data infrastructure for the province.
- Make the anonymized data available for educational research.

[2] We note that demographic data is often self-reported on a voluntary basis.

Introduction

Ontario's education sector is both data rich and information poor. Individual schools, colleges and universities each collect data, tied to an Ontario Education Number (OEN), about their students. The OEN effectively follows students along their educational pathways from kindergarten through to high school, where it is tied to characteristics including the courses they take, the grades they earn and the supports they access to help them succeed. The OEN is also used by Ontario's centralized college and university application centres to track applications and by publicly assisted colleges and universities to track, for example, the programs that individual students enrol in and their time to graduation.

Despite being ripe for analysis, data connected via the OEN is restricted to the Ontario government — and even then, OEN-linked data access is "limited and highly discretionary" (Gallagher-Mackay, 2017). The situation leaves educational institutions and analysts outside of government information poor — institutional data must be connected to understand how student experiences at any juncture (e.g., high school) affect their experiences and outcomes later (e.g., in college or the labour market). The lack of publicly available, connected educational data sources with information about individual students at different points in time makes the tasks of identifying problems and developing effective solutions complicated and murky, at best (Robson, 2021).

Despite this uncertainty, we do know that education systems advantage certain groups of students relative to others based on research from other jurisdictions and from Canada's Youth in Transition Survey.[3] Sociodemographic factors like household income, parental level of education, race, gender and high school achievement have all tended to impact whether a student pursues postsecondary, and the type of postsecondary education they choose (e.g., apprenticeship, college or university (Deller, Kaufman & Tamburri, 2019)). But in the absence of connected longitudinal data (that is, data that tracks the same individuals over time), it is impossible to know the extent to which the above or any other factors influence persistence and completion of postsecondary education. It is also difficult to confirm whether research findings from specific regions of the province, like Toronto, are broadly applicable.

The Hamilton Community Research Partnership (CRP) is a coalition of six organizations committed to supporting student achievement in Hamilton — and addressing gaps in

[3] An excerpt from a 2017 HEQCO report: "Statistics Canada's Youth in Transitions Survey (YITS) provided detailed longitudinal information about students' pathways, experiences and attitudes towards PSE. Combined with a substantial program-to-promote analysis of the rich data, YITS led to major findings on the determinants of PSE access, the impact of tuition fee changes, strong outcomes for immigrant students and many other topics. The last round of data collection for YITS was 2010" (Gallagher-Mackay, 2017).

available data about educational pathways — through a secure data-sharing network. It includes:

- two school boards: the Hamilton-Wentworth District School Board (HWDSB) and the Hamilton-Wentworth Catholic District School Board (HWCDSB);
- two postsecondary institutions: Mohawk College and McMaster University; and
- two public organizations focused on addressing systemic issues in education: the Hamilton Community Foundation (HCF) and the Higher Education Quality Council of Ontario (HEQCO).

The Hamilton CRP's primary goal is to break down information silos to better understand the education pathways of Hamilton students specifically. Despite steady gains in recent years, Hamilton high school students are graduating below the provincial average (McCullough, 2021) while the rate of adults without a postsecondary certificate, degree or diploma exceeds the provincial average (Statistics Canada, 2017). With access to longitudinal data, researchers, educators and policy-makers will be able to identify, examine and address systemic barriers affecting educational pathways.

To date, most research on educational pathways in Ontario comes from the Toronto District School Board (TDSB). Its data is unique in Canada; the board links demographic information from student census reports and other administrative data with key outcomes information, like high school graduation and postsecondary confirmations of acceptance (Gallagher-Mackay & Brown, 2021). In so doing, the TDSB has contributed to important research about predictors of, and barriers to, postsecondary access in Ontario. For instance, TDSB research has highlighted the predictive power of secondary grades and Grade 9 credit accumulation in relation to postsecondary access (Gallagher-Mackay & Brown, 2021; Robson et al., 2019; Brown & Tam, 2017). TDSB data has helped illuminate the harms inherent in Ontario's streaming practice, which separates students, based on perceived ability, into academic and non-academic courses in Grades 9 and 10 (Pichette et al., 2020).[4] The TDSB has also explored how student demographic characteristics, such as their race, sexual orientation and neighbourhood income, can affect postsecondary pathways (Robson et al., 2019; Brown & Tam, 2017). However, the TDSB being larger, more diverse and urban than other Canadian school boards raises questions about the broader applications of its research findings (TDSB, n.d.).

[4] Other boards, such as the Limestone District School Board in Kingston and the Durham District School Board, have contributed important research on the impacts of streaming by conducting pilot projects. Peel District School Board also conducted an important review with respect to systemic discrimination (Follwell & Andrey, 2021; Chadha, Herbert & Richard, 2020).

In connecting educational data in Hamilton for improvement purposes, the CRP also hoped to accomplish a second goal: demonstrating the untapped potential of data linkages using the OEN. The OEN offers the potential for depersonalized linkages of longitudinal education data for all students in the province; it was the key to pooling and connecting administrative data from Hamilton school boards, postsecondary institutions and application centres to create the CRP dataset. The Hamilton CRP is not the first initiative that thought to use creative data linkages to answer education-related questions in Ontario. The partnership comes on the heels of other data-linking initiatives involving the TDSB and application centres (see Robson et al., 2019); York University (see Parekh et al., 2020); and various datasets within Statistics Canada's Education and Labour Market Linkage Platform (see Walters et al., 2020; Brown et al., 2021). Outside of Ontario, British Columbia's Student Transitions Project, a joint initiative between B.C.'s education ministries and postsecondary institutions, links K–12 data to public postsecondary data to better understand student success (British Columbia, n.d.). Learning from these initiatives, the CRP sought to create the first community-level OEN-linked dataset that involves multiple school boards and multiple postsecondary partners, with each sharing data to better understand student pathways from secondary to and through both college and university in their community. The partners believe it should serve as encouragement, if not a proof of concept, for other communities across the province to follow suit. It should also serve as encouragement for the Ontario government to either use OEN-linked data to examine pathways in a similar manner or make this type of data available to organizations outside of government.

Finally, the partners sought to build trusting relationships and establish procedures that would enable ongoing, collaborative educational research in their community. This report, the first to draw from the CRP, is the result of the project's "pilot phase," designed to develop and test the data-sharing mechanism. The partners intend to expand the utility of the dataset gradually by adding cohorts and variables. The dataset will be used over the long term, in the absence of a provincial data infrastructure, to understand how Ontario's education system is advantaging or disadvantaging groups of students in Hamilton. It can also be used to evaluate the effectiveness of specific, local initiatives designed to help students succeed in K–12 and navigate pathways to and through postsecondary.

Research Questions and Methodology

In this first paper to draw from the Hamilton CRP dataset, we sought to test our data-sharing mechanism and answer a preliminary question: are OEN data linkages effective? If the answer proved to be yes, we would examine the educational pathways of Hamilton students from high school to and through postsecondary, answering the following questions: Who graduates high school in Hamilton? Who goes to college or university in

Ontario? And for those who study in Hamilton, who completes college or university programs? In other words, we wanted to understand the characteristics and experiences of those who navigate secondary and postsecondary in Hamilton: which students graduate high school, access PSE and graduate from their programs.

Variables of Interest

We looked to existing Canadian research to determine which variables we should include in our analysis (see Table 2 of **Appendix 1**). The de-identification process, which was essential given the pilot phase's focus on building trust and protecting privacy, resulted in a narrower set of independent variables than originally hoped. Even after we removed identifiers like student numbers and names, any characteristics that, in isolation or combination, could be used to re-identify an individual had to be generalized,[5] suppressed[6] or excluded from our analysis. The resulting dataset included: neighbourhood income, gender, first language, secondary grades, presence of an exceptionality, secondary stream, high school cumulative average, Grade 9 credit accumulation and postsecondary grade point average (GPA). Notably, our analysis is missing data pertaining to students' race, sexuality, Indigenous identity and parental education. And some of the variables we included lacked nuance. For example, the first language variable was reduced to English or 'other,' and we relied on neighbourhood rather than family income and the exceptionality variable aligned with the *Education Act*'s definition of the term, which includes students identified with giftedness. The lack of precision in these variables is an important limitation that the CRP partners plan to address in future iterations of the dataset.

Data Assembly and Analysis

Our analysis draws from the newly assembled CRP dataset, which links de-identified datasets from each educational partner using the OEN. McMaster, Mohawk, HWDSB and HWCDSB used a consistent formula to hash[7] the OENs in their datasets. This enabled data linkages and minimized the risk of re-identifying any students represented in the data. A designated Partnership Data Analyst (PDA), based at HWDSB, was the only person with access to the dataset once it was assembled. The PDA first merged the institutional datasets using the hashed OEN and probabilistic matching[8] based on variables like gender

[5] Generalization removes the granularity of data. For example, generalizing age could mean recoding someone aged 22.3 into a larger group of individuals aged 20–24. This makes it more difficult to reidentify an individual by replacing their numerical age with a range.
[6] Suppression is the removal of values from a dataset.
[7] Hashing is a process of mathematically obscuring an assigned number.
[8] We conducted probabilistic matching as a redundancy measure to ensure school board and postsecondary records with the same hashed identifier were referencing the same student. Probabilistic matching was

and calculated age. The PDA then removed the hashed OEN from the dataset and replaced it with a unique identifier. Aggregate results were sent to the other authors of this report using encrypted emails. Our companion report, *CRP Blueprint: How We Built a Community Data Infrastructure*, describes the dataset assembly in more detail, as well as rationale for each of the steps taken.

The PDA linked postsecondary application data from OCAS and the Ontario University Application Centre (OUAC) to school board data using the hashed OEN. This linkage permitted us to examine if Hamilton students accepted offers of admission to any Ontario college or university (not limited to Hamilton) and what programs they pursued. We note, however, that this step was complicated by the reality that some students in the sample applied to postsecondary programs in more than one application cycle. Our analysis focused on the earliest year that a student confirmed acceptance of an offer of admission to a postsecondary program.

Only the PDA had direct access to the CRP dataset once it was assembled. They conducted statistical analyses in STATA, including bivariate tests of association between the independent variables listed above and the following dependent variables: high school graduation,[9] confirmation of postsecondary admission, program choice, university graduation and college graduation. They also conducted multivariate analyses with logistic regressions to fit the predictor variables to the binary outcome of graduating secondary school or earning a postsecondary credential. The PDA ran multimodal logistic regressions to fit the predictor variables into the three possible outcomes of students confirming to a college, a university or neither postsecondary destination. They shared aggregate data with the other authors of this report using encrypted emails.

Sample Characteristics

The CRP dataset consists of data for 6,090 students who entered Grade 9 for the first time during the 2010-11 academic year. Most students (5,310) had sufficient data to be included in the study; we examined their educational trajectories from high school through to the end of their postsecondary studies.[10]

---

conducted in R with the *RecordLinkage* library. In the first iteration, the hashed identifier was blocked (i.e., identified as the variable across datasets for records to be compared) and linked on the variables of age (to one decimal place) and gender. This pass produced 2,519 matches out of 2,538 records. The 19 non-matching records were moved onto the second pass, which blocked again on the hashed identifier and linked on the postsecondary application numbers, producing 10 further matches. Following the two passes, we attained a match rate of 99.6%. We retained these students records for subsequent analyses.

[9] Within 10 years.

[10] Entry year to postsecondary was variable. Some students who graduated high school after seven years and then entered a four-year postsecondary program would not be shown as having graduated. The authors

Students in our sample had the option of taking academic, applied, open or locally developed courses in Grades 9 and 10. In senior grades, courses would have been geared towards postsecondary options: "university preparation," "university/college preparation," "college preparation," "workplace preparation" and "transfer" (Pichette et al., 2020).

The initial sample of high school students was evenly split in terms of gender (48% were female) and most (82%) spoke English as a first language. Approximately 13% of students were identified as having a learning exceptionality, including gifted. Just over two-thirds of students were in the academic stream, 28% were in majority applied courses and under 4% took mostly locally developed courses. The remaining 1.5% had undefined streams. Of these students, the average had earned 7.7 high school credits and had an average grade of 73%, but there was considerable range in both figures.

In terms of the neighbourhood household income categories, just over 23% of students were in the lowest neighbourhood income category ($30,000–$49,000) and similar percentages were in the middle two categories (33% and 36%). Around 8% were in the highest neighbourhood income category ($90,000 and up).

A subsample of 4,787 students applied to an Ontario college or university through a centralized application service and confirmed an offer of admission. About 41% of these students confirmed a university place, 40% confirmed a college place and 19% confirmed neither. For the subsample that went to postsecondary in Hamilton (2,321 students), roughly 67% graduated from university by November 2020 and around 57% graduated from college by December 2020.

For more detailed sample characteristics, see Table 3: Descriptive Statistics in **Appendix 1**.

Findings

All data tables are included in the **Appendices** of this report.

Are OEN Data Linkages Effective?

Data linkages proved to be an effective way of assembling a longitudinal dataset for education research; the CRP successfully linked administrative data from four institutions using the OEN and probabilistic matching. The resulting dataset had a 99.6% match rate, meaning we were able to connect school board and postsecondary data successfully for

acknowledge this was a limitation of this pilot (i.e., that a 10-year timespan is restrictive, especially when high school graduation rates are conventionally calculated within five years, and postsecondary graduation rates are conventionally calculated within seven years for university and three to seven years for college, depending on the program).

nearly all students who entered Grade 9 in 2010 in Hamilton and continued their postsecondary education at either of the local, publicly assisted postsecondary institutions, Mohawk College or McMaster University.

Are community-led OEN data linkages the *most* effective way of assembling a longitudinal dataset for education research? Unfortunately for the CRP members, the answer to that question is no. A provincial data infrastructure would provide more comprehensive data with which to explore the questions posed in this report, while being much more efficient. We explore these ideas in more detail in the Discussion section.

Who Graduated High School?

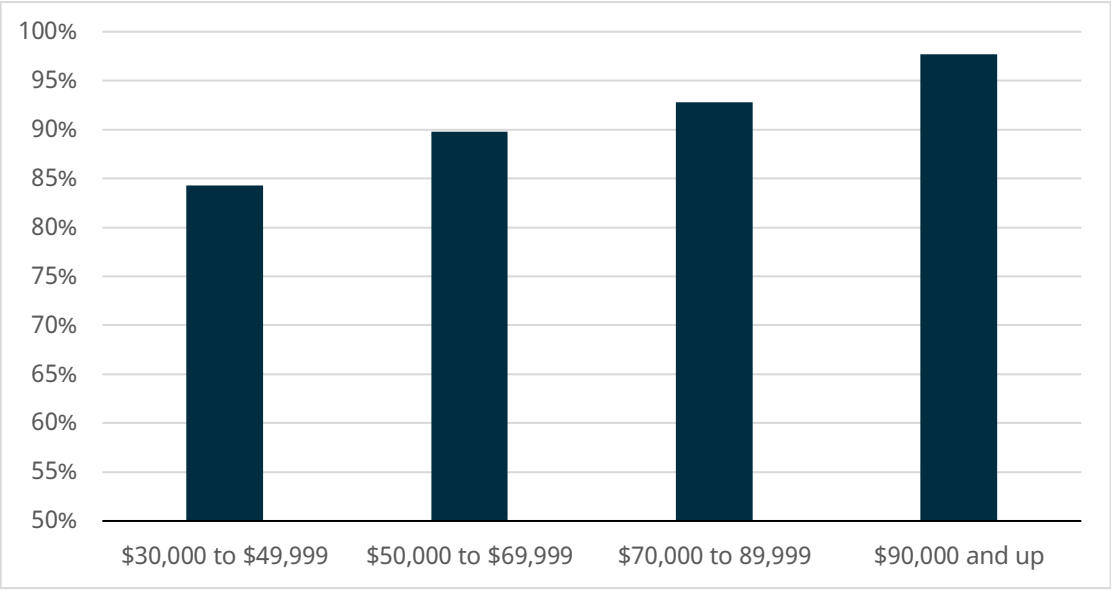
Most students (90%) in our sample graduated high school with an Ontario Secondary School Diploma (OSSD) within the 10-year timeframe we set for our analysis.[11] These students were predominantly in the academic stream; those concentrated in locally developed high school courses were around one-third as likely to graduate as those in the academic stream. We also saw a progressive increase in graduation rates by income quartile, as shown in Figure 1. Students who accumulated more Grade 9 credits and who had higher grades[12] were also more likely to graduate. Students with identified exceptionalities in our sample had significantly lower grade averages than those without. Just over 80% of students who had an exceptionality graduated high school, compared to just over 90% of students without an exceptionality.

[11] Our timeframe resulted in a higher graduation rate than what is typically reported in Hamilton, using a five-year timeframe. In 2020, the average five-year graduation rate across both Hamilton school boards was 84.5% (Ministry of Education, 2022)

[12] Higher secondary grades were associated more often with female students as well as with students with higher neighbourhood incomes.

Figure 1

High School Graduation by Neighbourhood Income Categories



Note. This figure shows the relationship between neighbourhood income and high school graduation rates. Graduation rates increase alongside increases in incomes.

Who Accessed Postsecondary?

Our linkage with application centre data (from OCAS and OUAC) allowed us to examine the characteristics of 4,787 students who applied and confirmed an offer to an Ontario college or university through a centralized application service. This group includes students who accepted offers to public Ontario postsecondary institutions outside of Hamilton. We could not account for those students who may have applied and confirmed offers of acceptance to institutions outside of Ontario (in other provinces or parts of the world).

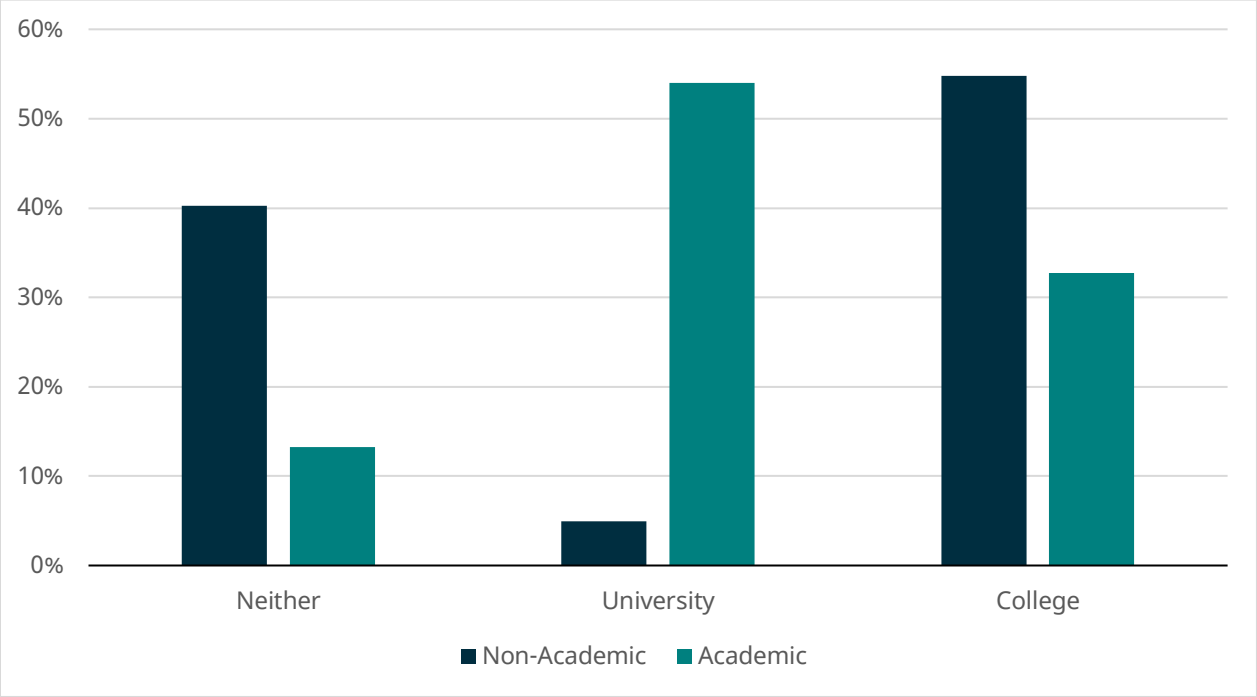
Approximately 41% of the sample confirmed an Ontario university offer of admission, 40% confirmed an Ontario college offer of admission and 19% confirmed neither (meaning they did not apply to an Ontario institution, or they applied and did not confirm an offer). On average, female-identifying students were more likely to confirm an offer of admission, and more specifically, confirm a university offer. Male-identifying students were more likely to confirm college offers.

Figure 2 illustrates that high school stream was an important factor affecting postsecondary access. Being in the academic stream was strongly associated with university confirmation. About 55% of students enrolled in the academic stream confirmed

a university offer in Ontario and another 33% confirmed a college offer. Only about 12% of students who had been in the academic stream did not confirm a postsecondary offer of admission. Meanwhile, 55% of students who had been in non-academic streams confirmed Ontario college offers, 5% confirmed university and 40% did not confirm an Ontario postsecondary offer. Students from the applied stream were 96% less likely to enrol in university relative to students in the academic stream.

Figure 2

Academic Stream by PSE Confirmation

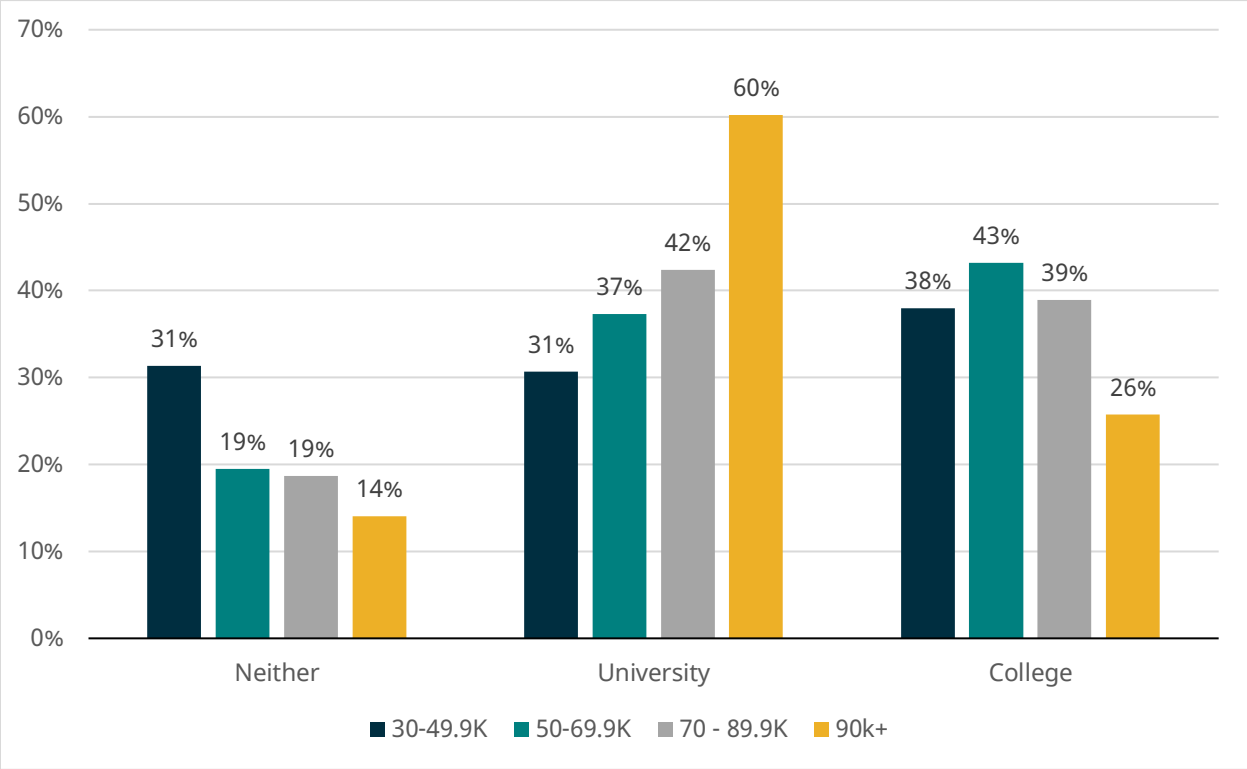


Note. This figure shows the relationship between high school stream (academic vs. non-academic) and PSE confirmation. Most students who confirmed university came from the academic stream.

Secondary grades and Grade 9 credits proved to be significant predictors of both college and university confirmations. We also found a positive association between neighbourhood income and confirmations of admissions to university, as illustrated in Figure 3. Regression analyses confirmed that income was an important predictor of university enrolment, with students in larger income bands having a greater likelihood of confirmation.

Figure 3

Neighbourhood Income Bands by PSE Confirmation



Note. This figure shows the relationship between students' neighbourhood income and postsecondary pathway. Students from higher-income neighbourhoods were most likely to pursue university.

What Programs Did Students Access?

College and university program codes were grouped into the two broad categories of STEM and BHASE,[13] aligned with Statistics Canada's classification structure displayed in Table 1.

[13] In some rare cases, we were unable to group the program code into one of these groups confidently — those programs were left as "indeterminate" in the analysis.

Table 1

STEM and BHASE Groupings

Grouping	Sub-groupings
STEM (Science, Technology, Engineering and Math)	• Science and science technology • Engineering and engineering technology • Mathematics and computer and information sciences
BHASE (Business, Humanities, Health, Arts, Social Science and Education)	• Business and administration • Arts and humanities • Social and behavioural sciences • Legal professions and studies • Healthcare • Education and teaching • Trades, services, natural resources and conservation

We found relationships between program choice and all independent variables included in this analysis, apart from exceptionality. Looking at college and university confirmations together, we found:

- Female-identifying students were far more likely to pursue BHASE programs compared to male-identifying students; there are nearly twice as many males as females in STEM programs.
- Students from academic high school streams were less likely to pursue BHASE programs; STEM programs were dominated by students who pursued the academic stream in high school.
- Higher secondary grades reduced the odds of students pursuing BHASE programs compared to STEM; students who pursued STEM programs tended to have considerably higher secondary averages (81%) than students who pursued BHASE programs (76%).

- Those in BHASE programs were less likely than students in STEM programs to be in the highest neighbourhood income band; more students who pursued STEM lived in neighbourhoods with average incomes of $90,000 or more.
- Those who spoke English as a first language were more likely to confirm BHASE programs; students who learned a language other than English first were slightly more represented in STEM than English native speakers. The reverse was true for BHASE.

Who Graduated from Postsecondary?

The CRP dataset allows us to look at graduation rates for the 2,321 students in our sample who pursued postsecondary in Hamilton at either McMaster University or Mohawk College. In the absence of a publicly available provincial data infrastructure, we cannot speak to the outcomes for those students who attended other postsecondary institutions in or outside of Ontario. Because our data is limited to a 10-year time span, it is possible that some students in our sample may have graduated college or university after spring 2021 or they may be poised to graduate soon — in either case, these students are not counted as having graduated from postsecondary in our analysis.

Working with the data available, we estimate that 67% of the students in our sample who attended university in Hamilton graduated as of November 2020, and about 57% of the students who attended college in Hamilton graduated as of December 2020. We note that we could not account for students who may have transferred out of either institution. Students entering either type of institution with higher secondary grades and more Grade 9 credits were more likely to graduate postsecondary (these were the same students who were more likely to graduate high school). Students with higher postsecondary GPAs also saw higher graduation rates than those who struggled academically; in fact, postsecondary grades were the most significant determinant of both university and college graduation. Postsecondary grades tended to be lower among male-identifying students, those who spoke a first language other than English, those who had an identified exceptionality and those coming from neighbourhoods with lower incomes. Postsecondary grades were also related to high school stream: students in the academic stream were more likely to have higher grades.

We also found an association between high school stream and postsecondary graduation. Students who were in the academic stream in high school graduated college and university in higher numbers. The streaming variable did not impact the PSE outcomes in our multivariate analysis. In other words, we did not see stream affect the likelihood of college or university graduation when controlling for other factors. This is likely a reflection of the streaming process itself. That is, nearly all university students and a

considerable proportion of college students in our sample came from the academic stream.

Neighbourhood income did not produce any significant results, nor did exceptionality, but this is likely because of the imprecise nature of these variables. We used large income bands to avoid identifying students, and the exceptionality category contained gifted students whose academic trajectories are typically quite different from students with other type of exceptionalities. The language variable appeared to have an effect only on college graduation, not university: native English speakers were 45% less likely to graduate college than those who learned another language first. We note that this finding is likely masking important differences that finer demographic details around immigrant-generation or ethno-linguistic background may reveal.

Discussion

The CRP succeeded in building a community-level dataset with information about students' educational journeys from secondary school through to college and university graduation — something unique in Ontario. Using the OEN to link school board and postsecondary data resulted in a match rate of very near to 100%, proving that educational data can be connected across multiple educational institutions in meaningful ways. Our companion report, *CRP Blueprint: How We Built a Community Data Infrastructure*, outlines steps that other communities can take to develop similar datasets with which to understand the advantages and disadvantages afforded to various groups of students.

The educational partners in the CRP collaborated with a goal of building trust and creating a culture where data sharing and analysis is welcome and encouraged. In the years to come, as the CRP and its stakeholders become increasingly comfortable using pseudonymized[14] data, we anticipate including additional and more nuanced demographic variables for analysis. We expect future phases of the CRP will parse specific languages spoken, family incomes and types of exceptionalities, such as giftedness, to capture the differential effects these categories may have on outcomes. And we are already one step closer to making that happen: as a result of participating in the CRP, each data-sharing partner has provided notice of collection to set the expectation among stakeholders that student data will be used and shared for ongoing improvement purposes. As the data quality improves, we will gain a more comprehensive look at systemic barriers affecting educational attainment as well as promising approaches to addressing those barriers. The requirement of all Ontario school boards to collect race-

[14] This is when direct identifiers such as names or student numbers are removed and replaced with something that is not identifying. Indirect identifiers like age and gender can remain in the dataset. Though unlikely, it is theoretically possible to identify pseudonymized individuals.

based data by January 2023 will also help to facilitate more nuance in years to come (Ontario Newsroom, 2020).

For now, this report contributes to a growing body of research into the factors affecting educational outcomes like high school graduation and postsecondary access. We found that established predictors, namely Grade 9 credit accumulation and secondary school grades, work well in Hamilton and may have broad applicability across Ontario. North American researchers have consistently found high school grade averages to be one of the most important predictors of participation in postsecondary education and persistence (Robson et al. 2019; Finnie et al., 2015; Hein et al., 2013; Dooley et al., 2012; Stewart et al., 2015). Postsecondary grades have been shown to carry a similar predictive power when it comes to student persistence (Stewart et al., 2015; Lopez-Rabson & McCloy, 2013). And Grade 9 credit accumulation is regularly cited as a key predictor of postsecondary participation (Robson et al., 2019; Finnie et al., 2015; Hein et al., 2013); data from the TDSB indicates students who complete fewer than eight Grade 9 credits in Grade 9 have limited chances of attending university (Robson et al., 2019). Accounting for other sociodemographic factors, our analysis confirmed that these attainment variables are the strongest determinants of graduation at both secondary and postsecondary levels for Hamilton students.

We found predictors of educational achievement to be closely tied to other demographic factors like neighbourhood income, secondary stream and exceptionality. In particular, and unsurprisingly, we saw income affording important advantages along educational pathways. Students from higher-income neighbourhoods were more likely to graduate high school, confirm a university offer and pursue STEM programs, all of which tend to result in higher incomes after graduation (Statistics Canada, 2017b; 2019). It would be interesting to explore the factors at play here: are higher-income students more likely to pursue tutoring for STEM outside of school, for example? This is an important area for future research.

Students who pursued the academic stream in high school gained similar advantages. Our analysis aligns with prior research out of the TDSB about the effects of streaming, suggesting students in non-academic courses are less likely to graduate high school or attend postsecondary than their peers in academic courses (Pichette et al., 2020). Academic and non-academic streams in Grades 9 and 10 lead to course options aligned with specific post-graduation pathways in Grades 11 and 12 (e.g., “university preparation” and “college preparation”). Despite an intended connection between the applied stream and college preparation courses, and eventually, college enrolment, we found students in the applied stream were less likely to confirm a postsecondary offer of admission, including to college, relative to students in the academic stream. About 88% of students who had been in the academic stream confirmed a postsecondary offer, while only about

60% of students in non-academic streams did so. We also found an association between high school stream and postsecondary graduation. Students who were in the academic stream in high school graduated both college and university in higher numbers.

Research out of the TDSB and Peel District School Board indicate an overrepresentation of racialized students in non-academic streams (Follwell & Andrey, 2021; Pichette et al. 2020), meaning equity-deserving students have historically been most disadvantaged by this system. The Government of Ontario made a progressive decision to de-stream Grade 9 math in 2020-21 and has announced plans to de-stream all Grade 9 courses beginning in 2021-22 (Ontario Newsroom, 2021). Ongoing longitudinal data collection and analysis is needed to ensure these policies are implemented as planned and result in equity improvements. This kind of analysis should also inform related policy frameworks, which should include a long-term strategy for de-streaming compulsory Grade 10 courses.

While community-level coalitions like the CRP can generate research that informs these sorts of policy decisions and evaluations, a provincial data infrastructure would be far more straightforward and informative. The CRP dataset is limited to a 10-year timeframe and can explore pathways only for the students who remain in Hamilton for their postsecondary studies. A provincial data infrastructure would allow researchers to explore pathways for a broader group of students, including those who enter the workforce before pursuing postsecondary and those who pursue programs outside of the community where they attended high school. It would also allow for analysis of barriers as they emerge (rather than drawing from historic data, as we have). Lastly, by virtue of being larger, a provincial data infrastructure would enable the inclusion of more precise demographic details. The B.C. Student Transitions Project is evidence that this kind of infrastructure can be built and maintained in a way that informs policy while protecting student privacy.

Conclusion and Recommendations

In a world where data is ubiquitous, information should not be so hard to access. Through our collaboration, the CRP built a connected dataset to answer important equity questions in the absence of anything similar available at the provincial level. A provincial-level data infrastructure would allow us to answer the research questions addressed in this report with more completeness and certainty.

With this in mind, we offer the following recommendations to the Ontario government, with the ultimate goal of improving educational equity in Ontario:

- Require that institutions collect consistent administrative and demographic data.

    - School boards, application centres, colleges and universities should all be required to collect demographic information (e.g., about race, Indigenous identity, disability, sexual orientation/identification, parental education and immigrant/visible minority status), both regularly and consistently. Ontario's *Anti-Racism Act, 2017* and Anti-Racism Data Standards represent an important step in this regard.[15]
    - Demographic and administrative data (including courses taken/completed, grades and graduation) should be connected to the OEN, which will in turn enable linkages with labour market datasets such as T1 Family Files (T1FF).
    - Ontario should facilitate consistency in the creation of data around key markers like programs codes/categories (see Appendix 2), working with application centres and postsecondary institutions.
    - In addition to feeding into a provincial data infrastructure, consistent data collection will position institutions to collaborate in evaluating cross-institutional initiatives.
- Build a provincial educational data infrastructure.
    - Use the OEN to create a longitudinal dataset that tracks students from the time they enter primary school through to the workforce and back again, so we can understand lifelong learning journeys and barriers along the way (for equity-seeking students in particular).
    - This would also enable research about Ontarians who pursue pathways other than college or university after high school, including apprenticeship training, community living and the workplace.
- Make the anonymized data available for educational research.
    - Data should be available for researchers to evaluate specific policies and initiatives, such as the de-streaming policy, and to inform both system-wide as well as more localized improvements.

We believe this project serves as proof of concept to the Ontario government, demonstrating the power of connected data. In the meantime, Ontario communities should consider building similar partnerships to explore these questions. Our companion report, *CRP Blueprint: How We Built a Community Data Infrastructure*, which summarizes the steps we took to assemble the CRP dataset, is available as a resource.

[15] In accordance with the *Anti-Racism Act* and Anti-Racism Data Standards, all school boards in Ontario will be required to collect race-based data by January 1, 2023 (Ontario Newsroom, 2020).

References

British Columbia, (n.d.). Student Transitions Project https://www2.gov.bc.ca/gov/content/education-training/post-secondary-education/data-research/student-transitions-project#:~:text=The%20Student%20Transitions%20Project%20uses,post%2Dsecondary%20education%20and%20graduate.

Brown, L. (2014, October 6). Gay students more likely to drop out, TDSB survey learns. *Toronto Star*. https://www.thestar.com/yourtoronto/education/2014/10/06/gay_students_more_likely_to_drop_out_tdsb_survey_learns.html

Brown, R. S. & Tam, G. (2017). *Grade 9 cohort post-secondary pathways, 2011-16. Fact sheet 3, November 2017.* Toronto: Toronto District School Board. https://www.tdsb.on.ca/Portals/research/docs/reports/FS3%20Grade%209%20Cohort%20Post-Sec%20Pathways%202011-16%20FINAL.pdf

Brown, R., Walters, D., Parekh, G.; & Conley, C. (2021). *Overall Patterns: Match of the TDSB 2004-2006 Grade 9 Cohorts with PSIS Data*. Toronto: Toronto District School Board. https://cpb-ca-c1.wpmucdn.com/sites.uoguelph.ca/dist/1/313/files/2021/12/Descriptive-Report-Overall-Patterns-TDSB-Match-to-Postsecondary-AODA-Compliant.pdf

Chadha, E., Herbert, S. & Richard, S. (2020). *Review of the Peel District School Board*. Toronto: Ministry of Education. http://www.edu.gov.on.ca/eng/new/review-peel-district-school-board-report-en.pdf

Chatoor, K., MacKay, E. & Hudak, L. (2019). *Parental Education and Postsecondary Attainment: Does the Apple Fall Far from the Tree?* Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/pub/parental-education-and-postsecondary-attainment-does-the-apple-fall-far-from-the-tree/

Chatoor, K. (2021) *Postsecondary Credential Attainment and Labour Market Outcomes for Ontario Students with Disabilities*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2021/06/Postsecondary-Credential-Attainment-and-Labour-Market-Outcomes-for-Ontario-Students-with-Disabilities_FINAL.pdf

Chatoor, K., Courts, R., Han, J., Barclay, V., & Colyar, J. (2022) *Access Programs in Ontario: OPAIP and Pathways to Education*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2022/01/Access-Programs-in-Ontario_FINAL1.pdf

Deller, F., Kaufman, A. & Tamburri, R. (2019*). Redefining Access to Postsecondary Education*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2020/02/Formatted-Access-Paper.pdf

Dooley, M. D., Payne, A. A., & Robb, A. L. (2012). Persistence and Academic Success in University. *Canadian Public Policy, 38*(3), 315–339. Doi:10.1353/cpp.2012.0028.

Ferguson, S. (2016). Women and Education: Qualifications, Skills and Technology. *Women in Canada: A Gender-based Statistical Report*. Ottawa: Statistics Canada. https://www150.statcan.gc.ca/n1/en/pub/89-503-x/2015001/article/14640-eng.pdf

Finnie, R., Pavlic, D., Jevtovic, N., & Childs, S. (2015). *Post-Schooling Outcomes of University Graduates: A Tax Data Linkage Approach*. Ottawa: Education Policy Research Initiative. https://ruor.uottawa.ca/handle/10393/32088

Follwell, T. & Andrey, S. (2021). *How to End Streaming in Ontario Schools*. Toronto: Ontario 360. https://on360.ca/policy-papers/how-to-end-streaming-in-ontario-schools/

Ford, R., Hui, T. S., & Nguyen, C. (2019). *Postsecondary Participation and Household Income*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2020/02/Formatted-SRDC-PSE-Access-and-Income-1.pdf

Gallagher-Mackay, K. (2017). *Data Infrastructure for Studying Equity of Access to Postsecondary Education in Ontario*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2020/03/FINAL-Data-Infrastructure.pdf

Gallagher-Mackay, K. & Brown, R. S. (2021) *The Impact of School Closures and Emergency Remote Learning on Postsecondary Transitions in 2020/21: Findings from Toronto*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2021/11/FORMATTED-Covid-in-TDSB-postsecondary-transitions.pdf

Hein, V., Smerdon, B., & Sambolt, M. (2013). *Predictors of Postsecondary Success*. Washington: College and Career Readiness and Success Center (American Institutes for Research). https://files.eric.ed.gov/fulltext/ED555671.pdf

Lopez-Rabson, T. S. & McCloy, U. (2013). *Understanding Student Attrition in the Six Greater Toronto Area (GTA) Colleges*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2020/03/Final-Formatted-Seneca-Student-Attrition.pdf

McCullough, K. (2021, March 13). Hamilton public school board graduation rates increase slightly but continue to fall short of Ontario average. *The Hamilton Spectator*.

https://www.thespec.com/news/hamilton-region/2021/03/13/hwdsb-school-graduation-rates.html

Ministry of Education. (2022). *School Board Progress Reports*. Toronto: Ministry of Education. https://www.app.edu.gov.on.ca/eng/bpr/allBoards.asp?chosenIndicator=11

Nunez, A., Rios-Aguilar, C., Kanno, Y., & Flores, S. (2016). Chapter 2: English Learners and Their Transition to Postsecondary Education. *Higher Education: Handbook of Theory and Research 31* (pp. 41–90). Springer International Publishing Switzerland. DOI 10.1007/978-3-319-26829-3_2 https ://www.bu.edu/wheelock/files/2016/10/Nunez-et-al.-2016.pdf

Ontario Ministry of Education. (2017). Special Education in Ontario: Kindergarten to Grade 12. http://www.edu.gov.on.ca/eng/document/policy/os/2017/SpecEdFinal2018.pdf

Ontario Newsroom. (2020, July 9). *Ontario Taking Bold Action to Address Racism and Inequity in Schools* [Press release]. Toronto: Ontario Government. https://news.ontario.ca/en/release/57543/ontario-taking-bold-action-to-address-racism-and-inequity-in-schools

Ontario Newsroom. (2021, June 9). *Modernized Math Course Prepares Students for Jobs of the Future and Life Skills* [Press release]. Toronto: Ontario Government. https://news.ontario.ca/en/release/1000301/modernized-math-course-prepares-students-for-jobs-of-the-future-and-life-skills

Ontario Newsroom. (2022, February 25). *Ontario Supporting the Success of Black Students* [Press release]. Toronto: Ontario Government. https://news.ontario.ca/en/release/1001660/ontario-supporting-the-success-of-black-students

Parekh, G., Brown, R. S., & James, C. E. (2020). *Who comes to York? Access, Participation, and Graduation Trends*. Toronto: York University. https://www.researchgate.net/publication/356392516_Who_Comes_to_York_Access_Participation_and_Graduation_Trends

Pichette, J., Deller, F., & Colyar, J. (2020) *Destreaming in Ontario: History, Evidence and Educator Reflections*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2020/10/Destreaming-in-Ontario_FORMATTED.pdf

Robson, K., Maier, R., Anisef, P., & Brown, R. S. (2019). *High School Success and Access to Postsecondary Education*. Toronto: Higher Education Quality Council of Ontario. https://heqco.ca/wp-content/uploads/2020/02/Formatted-CRP-report_FINAL.pdf

Robson, K. (2021). An Essay on the Challenges of Doing Education Research in Canada. *Journal of Applied Social Science, 15*(2), 183-196. doi:10.1177/19367244211003471 https://journals.sagepub.com/doi/full/10.1177/19367244211003471

Statistics Canada. (2016). *Variant of CIP 2016 – STEM and BHASE groupings.* Statistical classifications. Ottawa: Statistics Canada. https://www23.statcan.gc.ca/imdb/p3VD.pl?Function=getVD&TVD=401856

Statistics Canada. (2017). *Hamilton [Census metropolitan area], Ontario and Ontario [Province]* (table). Census Profile. 2016 Census. Statistics Canada Catalogue no. 98-316-X2016001. Ottawa: Statistics Canada.

Statistics Canada. (2017b). Census in Brief: Is field of study a factor in the earnings of young bachelor's degree holders? Retrieved from https://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016023/98-200-x2016023-eng.cfm

Statistics Canada. (2019). Average earnings or employment income, by age group and highest certificate, diploma or degree. Retrieved from https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3710015201

Stewart, S., Hun Lim, D., & Kim, J. (2015). Factors Influencing College Persistence for First-Time Students. *Journal of Developmental Education, 38*(3), 12–20. https://files.eric.ed.gov/fulltext/EJ1092649.pdf

Sweet, R., Pullman, A., Adamuti-Trache, M., & Robson, K. (2019). Ethno-linguistic Patterns of Degree Completion in BC Universities: How important are high-school academic achievement and institution of entry? *The Canadian Journal of Higher Education, 49*(3), 69–87. https://www.erudit.org/en/journals/cjhe/1900-v1-n1-cjhe05066/1066636ar.pdf

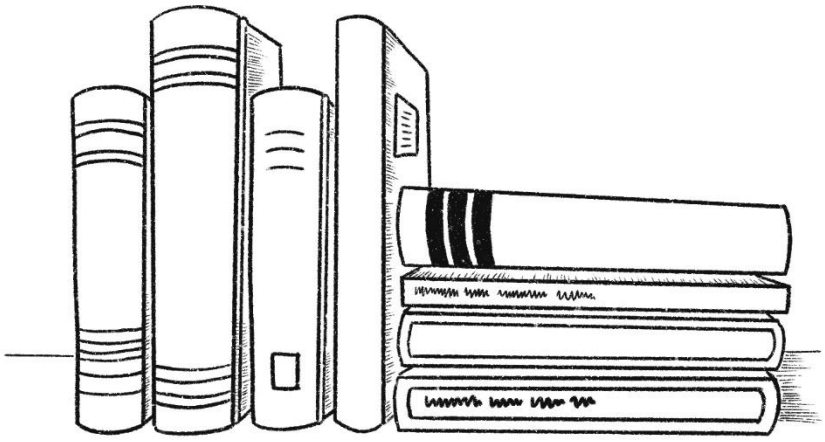
Toronto District School Board. (n.d.). About Us. Toronto: Toronto District School Board. https://www.tdsb.on.ca/About-Us

Walters, D., Brown, R., Parekh, G., Einmann, T. & Bader, D. (2020). *Student Loan Outcomes of Ontario Transfer Students*. Toronto: Ontario Council on Articulation and Transfer.

Walton, P., Hamilton, K., Clark, N., Pidgeon, M., Arnouse, M. (2020). Indigenous University Student Persistence: Supports, Obstacles, and Recommendations. *Canadian Journal of Education, 43*(2), 431–464.

Yau, M., Rosolen, L., & Archer, B. (2015). *Census portraits, understanding our students' backgrounds: sexual orientation report* (Report No. 14/15–23). Toronto: Toronto District School Board.

https://www.tdsb.on.ca/Portals/research/docs/reports/Portrait_Census2011-12_SexualOrient_FINAL_report.pdf



The Power of Connected Data: Charting Student Pathways to and through Postsecondary in Hamilton

Appendices

Appendix 1: Data Tables

Table 2

Variables of Interest

Variable of Interest	Sample of Existing Research
Income	A recent HEQCO study drawing on Toronto District School Board (TDSB) data found students from neighbourhoods with higher incomes were more likely to be accepted to university (Robson et al., 2019). Another HEQCO study that drew on census data found that, across Canada, postsecondary participation rates tend to increase with family income, and this pattern is mainly driven by university participation (Ford et al., 2019).
Gender	Canadian women are shown to accept, attend and graduate university at higher rates than men (Ferguson, 2016).
Sexuality	A recent HEQCO report acknowledged a "a near complete absence of reliable data for LGBTQ+ students" (Chatoor et al., 2022). Limited data from the TDSB suggests students who identify as lesbian, gay, bisexual or transgender are less likely to graduate from high school (Brown, 2014). Another recent TDSB study found LGBTQ students were less likely than heterosexual students to be on track to graduate high school at the end of Grade 10 (Yau et al., 2015).
First language	Researchers in British Columbia found considerable variation in bachelor's degree completion across ethno-linguistic groups. Korean and Vietnamese/Tagalog speakers, for example, were shown to consistently lag the predicted performance of other ethno-linguistic groups (Sweet et al., 2019). American research indicates students with limited English proficiency tend to "complete high school, enrol in college, and graduate from college at far lower rates than their peers" (Nunez et al., 2016). Of course, students whose first language is not English may still be proficient in the language.
Exceptionality/Disability	In the K-12 system, having an identified exceptionality can qualify a student for special education programs and

	services. There are five categories of exceptionalities in Ontario: behavioural, communicational, intellectual, physical and multiple (Ontario Ministry of Education, 2017). Notably, students identified with giftedness (i.e., advanced general intellectual ability) are included in the categorization of intellectual exceptionalities. Previous HEQCO reporting suggests having an exceptionality aside from giftedness lowers the odds of confirming postsecondary acceptance (Robson et al., 2019). In postsecondary, students who self-identify as having a disability as defined by the *Accessibility for Ontarians with Disabilities Act* can register with their institution to access accommodations. A recent HEQCO report found that, generally, individuals with disabilities have a lower rate of overall postsecondary attainment (including college, bachelor's and post-bachelor's credentials) than those without disabilities. More specifically, students with learning and physical disabilities in Ontario are the least likely to acquire a postsecondary credential, and especially a bachelor's or post-bachelor's credential (Chatoor, 2021).
First-generation status	First-generation students come from families where parents or guardians did not complete a PSE credential. While there is very limited data available about this population of students, previous HEQCO reports have highlighted an attainment gap between high school students from families in which no parent participated in postsecondary education and students whose parents have credentials; Deller et al. (2019) noted that when students do participate, first-generation youth are more likely to pursue two-year college programs rather than four-year university degrees. Estimates from Statistics Canada's Longitudinal International Survey of Adults linked to the Intergenerational Family File show parental education is a major determinant of postsecondary participation (Chatoor et al., 2019).
Indigeneity	In Canada, the grouping of "Indigenous people" includes First Nations, Inuit and Métis peoples, though it is important to note there is a diversity in the histories and cultures across these groupings.

	Data from Statistic's Canada's Youth in Transition Survey (YITS), which ended in June 2010, indicated Indigenous students in Canada have had significantly lower postsecondary participation rates and higher dropout rates than non-Indigenous students. YITS data indicates that at the time of the last survey, 9.8% of Indigenous respondents completed a university degree, compared to 26.5% of the general population (Walton et al., 2020). A recent HEQCO report drew on the 2016 Canadian Census for estimates and found that a larger proportion of Indigenous individuals have no postsecondary credential than non-Indigenous individuals in Ontario. The same report found Indigenous students account for 3.7% of college students and 1.9% of university students (Chatoor et al., 2022).
Race/Visible minority	HEQCO's *Redefining Access to Postsecondary Education* paper acknowledged that "students from racialized backgrounds, especially boys, are more likely to be expelled and suspended while in high school and streamed into vocational programs" which are highly determinant of who goes to postsecondary (Deller et al., 2019). Since then, a HEQCO report examined enrolment patterns using 2016 Census data. It found that "34% of adult visible minorities in Ontario have no postsecondary credential compared to 38% of Ontario adults who are not visible minorities." The same report noted a great deal of variation within the grouping of "visible minority." For example, "for 44% of Koreans, 15% of Black Canadians, and 29% of South Asians, the highest credential obtained is a bachelor's degree. However, for 42% of Latin Americans, 24% of Filipinos, and 58% of Southeast Asians, the highest credential earned is a high school diploma" (Chatoor et al., 2022).
Secondary stream	As explained in a recent HEQCO report, from 1999 to 2021, Ontario students in Grades 9 and 10 have had to select between academic, applied, locally developed and open-level courses. Students who have chosen applied courses are less likely to graduate high school or attend postsecondary than their peers in academic courses. TDSB data shows that "racialized students, especially Black males and students from lower-income families, are more likely to enrol in

	applied courses, and students from wealthier families are more likely to enrol in the academic stream" (Pichette et al., 2020). Recognizing that the practice of streaming has historically disadvantaged racialized students, the Ontario government announced a plan to de-stream Grade 9 math beginning in September 2021 (Ontario Newsroom, 2020). All remaining Grade 9 courses will be de-streamed in September 2022 Ontario Newsroom, 2022). Grade 10 courses remain streamed for the time being.
Secondary grades	North American researchers have consistently found high school grade averages to be one of the most important predictors of participation in postsecondary education (Robson et. al. 2019, Finnie et. al., 2015, Hein et al., 2013). High school grades are also shown to be predictive of postsecondary persistence (Dooley et al., 2012; Stewart et al., 2015)
Grade 9 credits	Like secondary grades, Grade 9 credit accumulation is often cited as a key predictor of postsecondary participation (Robson et al., 2019; Finnie et al., 2015; Hein et al., 2013). A recent HEQCO report explains that "generally, students who complete fewer than eight credits in Grade 9 have a quite limited chance of going to university" (Robson et al., 2019).
Postsecondary Grade Point Average (GPA)	First semester GPA is shown to be a significant predictor of student persistence in North America (Stewart et al., 2015). A 2013 HEQCO study found students with lower GPAs were more likely to leave college without completing their program (Lopez-Rabson & McCloy, 2013).

Table 3

Descriptive Statistics

Variable	N	Mean	St. Dev	Min	Max
Graduated High School	5,310	0.903		0	1
Female	5,310	0.479		0	1

Stream					
Academic	5,310	0.665		0	1
Applied	5,310	0.282		0	1
Locally Developed	5,310	0.036		0	1
Undefined	5,310	0.017		0	1
First Language English	5,310	0.822		0	1
Neighbourhood Income					
$30,000 to $49,999	5,310	0.232		0	1
$50,00 to $69,999	5,310	0.328		0	1
$70,000 to $89,999	5,310	0.358		0	1
$90,000 and up	5,310	0.082		0	1
Exceptionality	5,310	0.133		0	1
Grade 9 Credits	5,310	7.760	1.188	0	12
Secondary Grades	5,310	73.033	12.327	11.76	98.300

Table 4

Bivariate Tests of Association Between Independent and Dependent Variables

Variables	High School Grades
Female	*t*= -14.282***
Stream	*F*= 765.05 ***
First Language English	*t*= 1.659
Neighbourhood Income	*F*= 92.52 ***
Exceptionality	*t*=12.718***
Grade 9 Credits	ρ= 0.505***
Secondary Grades	

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Female students had secondary averages that were five percentage points higher than their male counterparts. While first language did have a marginal statistical significance (p<0.10), the differences between the two groups were less than two percentage points. In contrast, we observed a large difference in average high school grades between students with exceptionalities and those without, with a gap of around six percentage points.

Table 5

Ordinary Least Squares Regressions Predicting Secondary Achievement: Unstandardized Coefficients

	HS Grades	HS Grades
Female	3.269***	3.310***
Stream		
Applied	-13.15***	-10.09***
Locally Developed	-14.88***	-8.929***
Undefined	-15.51***	-11.40***
First Language English	-0.933*	-1.128***
Neighbourhood Income		
$50,000 to $69,999	1.874***	1.311***
$70,000 to $89,999	2.754***	2.058***
$90,000 and up	5.548***	4.378***
Exceptionality	0.969*	0.377
Grade 9 Credits		4.113***
Constant	74.56***	42.24***
R^2	0.330	0.468
N	5310	5310

Table 5 presents the ordinary least squares regression predicting secondary grades. Variables are added in steps to account for the know association between prior attainment and current attainment. In the first model, all variables are statistically significant. Females have about a 3% higher average grade, controlling for the other variables in the model. In terms of streams, compared to the academic stream (an omitted category), all other streams have strong negative coefficients. Native English speakers also had a significantly negative coefficient — something that differs from the bivariate models we examined earlier. Neighbourhood income was associated with grades in the expected manner, with average grade rising significantly as the neighbourhood income increased. In the first model, having an exceptionality is positively associated with grades; however, when we add the control for Grade 9 credit accumulation, this is the only variable that drops from statistical significance, indicating the effect of exceptionality on grades is mediated through prior attainment.

Table 6

Bivariate Tests of Association Between Independent and High School Graduation

	Graduated High School
Female	z=-4.599***
Stream	$\chi^{2=}$514.065***
First Language English	z=0.224
Neighbourhood Income	$\chi^{2=}$92.4592***
Exceptionality	z=8.385***
Grade 9 Credits	t = 31.641***
Secondary Grades	t = 46.661***
PSE Grades	

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6 presents the bivariate associations between high school graduation and various associated characteristics. The tests of association suggest that sex, academic stream, neighbourhood income, exceptionality and high school academic achievements are significantly associated with high school graduation.

Table 7

Logistic Regressions Predicting High School Graduation Controlling for Sex, Stream, First Language, Neighbourhood Income, and Educational Achievement: Odds Ratios

	Graduated High School
Female (1=yes)	0.770*
Stream (reference=academic)	
Applied	0.994
Locally Developed	0.347***
Undefined	0.802
First Language English (1=yes)	1.015
Neighbourhood Income (reference=$30,000 to $49,999)	
$50,00 to $69,999	0.891
$70,000 to $89,999	0.920
$90,000 and up	1.648
Exceptionality (1=yes)	1.044
Grade 9 Credit Accumulation	0.975
Secondary Average	1.209***
N	5310
Pseudo R2	0.4593

Exponentiated coefficients

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7 presents a multivariate estimation of high school. Sex was statistically significant for high school graduation, and in the opposite direction to what is typically seen the literature. In other words, being female reduced the odds of graduating high school by 23% compared to males. Compared to academic streams, those in locally developed courses were 65% less likely to graduate high school. The only other significant predictor

of high school graduation was high school grades; these were by far the most important predictor of high school graduation when further analyses to examine the importance of factors was carried out (not shown). In general, this model is not particularly informative, which is possibly due to several important sociodemographic factors not being available in these data.

Table 8

Descriptive Statistics for Current Subsample (N=4787)

	Mean	S.D	Min	Max
Confirmed University	0.406		0	1
Confirmed College	0.403		0	1
Confirmed Neither	0.190		0	1
Female (1=yes)	0.489		0	1
Other Stream	0.289		0	1
Academic Stream	0.711		0	1
First Language English	0.821		0	1
Neighbourhood Income				
$30,000 to $49,999	0.217		0	1
$50,000 to $69,999	0.326		0	1
$70,000 to 89,999	0.368		0	1
$90,000 and up	0.089		0	1
Has Exceptionality	0.121		0	1
Grade 9 Credits	7.915	0.900	0	12
High School Grades	75.204	10.117	34.1	98.3

Table 8

*Descriptive Statistics for Current Subsample (N=4787)*Table 8 presents the descriptive statistics for the subsample that includes information on confirmation to Ontario postsecondary institutions. We can see that around 41% of the sample confirmed a

university place, 40% confirmed a college place and 19% confirmed neither. It should be noted that we use a binary streaming variable in the forthcoming analyses where 1=academic high school stream and 0=other streams; there were not enough cases in the locally developed and undefined to allow separate categories, particularly since the vast majority of university-confirmed students had an academic pathway.

Table 9

Bivariate Tests of Association Between Independent and Dependent Variables

	Confirmed Neither	Confirmed University	Confirmed College
Female	*z*= 6.787***	*z* = -8.512***	*z* =2.8585**
Academic Stream	*z* = 20.819***	*z* = -31.603***	*z* = 14.284***
First Language English	*z* = -6.112***	*z* = 5.801***	*z* = -0.709***
Neighbourhood Income	*F*= 29.50 0***	*F*=40.850***	*F*= 14.620***
Exceptionality	*z* = -9.048***	*z* = 11.438***	*z* = -3.905***
Grade 9 Credits	*F*= 631.000***	*F*= 14.440***	*F*= 3.610***
Secondary Grades	*F*= 10.970***	*F*= 2596.900***	*F*= 508.250***

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9 presents the binary associations between the postsecondary confirmation pathways and the associated factors of interest. In these exploratory analyses, all factors were statistically significant.

Table 10

Multinomial Logistic Regression of Confirming an Offer to Ontario Postsecondary (reference=confirmed at neither): Odds Ratios

	Confirmed University	Confirmed University	Confirmed University	Confirmed College	Confirmed College	Confirmed College
Female	1.681***	1.690***	1.042	1.243**	1.251**	1.179*
Stream				1	1	1

Academic	29.16***	24.15***	9.823***	1.694***	1.586***	1.433***
First Language English	0.381***	0.376***	0.355***	0.594***	0.573***	0.579***
Neighbour-hood Income				1	1	1
$50,000 to $69,999	1.676***	1.743***	1.638***	1.786***	1.762***	1.721***
$70,000 to $89,999	1.813***	1.866***	1.590***	1.650***	1.618***	1.572***
$90,000 and up	3.491***	3.454***	2.710***	1.476*	1.428	1.367
Exceptionality	0.812	0.768	0.623**	0.843	0.847	0.833
Grade 9 Credits		1.815***	1.131		1.181***	1.118**
Secondary Grades			1.197***			1.020***
Pseudo R2				0.1423	0.1512	0.2593
N				4787	4787	4787

The variables in Table 10 were added in stages so that academic achievement measures were included in the second and third steps. In terms of predicting university confirmation, females were over 1.5 times more likely to enrol than males, but when secondary grades were accounted for in the third model, this effect was no longer significant. A similar effect was found in terms of college enrolment, although the effect was slightly smaller.

Students from the applied stream were 96% less likely to enrol in university (relative to students in the academic stream), and this remained significant, although weakened once prior achievement was factored in. Applied students were also less likely than academic stream students to enrol in college, but this effect was much smaller than in university enrolment and also weakened when grades were factored. Speaking English as a first language also reduced the odds of both university and college enrolment, although the

effect was larger in the case of university enrolment. This effect remained significant regardless of grades. Income was found to be an important predictor of university enrolment, with larger income bands having a greater likelihood of enrolment in university in particular. The effect was also seen in college enrolment, with the exception of the highest income band. Exceptionality was not a significant predictor for either enrolment pathway when accounting for other variables in the model. Looking to the model fit, we also observe a much greater improvement in fit once prior attainment (secondary grades) are accounted for.

Table 11

Bivariate Associations Between Confirmed Postsecondary Program Type and Independent Variables of Interest

	BHASE	STEM	Indeterminate/Other
Female	*z* = -10.542***	*z* = 12.088***	*z*= -0.5005
Stream	*z* = 5.029***	*z* = -8.933***	*z* = 4.480***
First Language English	*z* = -1.709*	*z* = 5.173***	*z* = -4.388***
Neighbourhood Income	*F*= 1.000	*F*=3.763***	*F*= 2.49***
Exceptionality	*z* = -0.418	*z* = 0.4450	*z* = 0.026
Grade 9 Credits	*F*=1.63*	*F*=3.69***	*F*= 2.03***
Secondary Grades	*F*= 83.75***	*F*= 287.26***	*F*= 72.12***

Table 11 reveals statistically significant associations between program choice and almost all our independent variables of interest apart from exceptionality.

Table 12

Bivariate Tests of Association Between Independent and Dependent Variables

	Graduated McMaster University	Graduated Mohawk College
Female	*z*=-3.357***	*z*=-4.971***
Stream	*z = -15.232***	*z = 6.4814***

First Language English	z=-0.571	z=0.034
Neighbourhood Income	χ^2=2.857	$\chi^{2=}$2.735
Exceptionality	z=0.3858	z=0.0451
Grade 9 Credits	t = 1.688*	t = 5.485***
Secondary Grades	t = 8.671***	t = 14.592***
PSE Grades	t = 14.939***	t = 30.462***

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12 presents results from the test of significance, which was determined by the level of measurement of the independent variable: bivariate nominal tests used a test of proportions, bivariate tests of continuous measures used two-sample t-tests, and tests where the independent variable of interest was multinomial used a Chi-Squared test of independence.

Appendix 2: Challenges with Program Codes

Tied to the question of who goes to postsecondary, the CRP sought to better understand factors affecting program choice. This was not a straightforward topic to explore in large part due to challenges with the classification of program codes.

Statistics Canada offers a classification structure which includes three STEM categories and seven BHASE categories (Statistics Canada, 2016). This would be an excellent structure for sorting and analyzing data about program choice were it not for one problem: postsecondary program codes in Ontario are not standardized across institutions. Each Ontario postsecondary institution has the discretion to categorize their program codes as they see fit, and there is no key to map these codes to Statistics Canada's Classification of Instructional Program (CIP) codes.

To work with CIP codes, we first had to combine some of the categories to ensure the information was not identifying, which resulted in seven generalized CIP code categories: two STEM categories and five BHASE categories. Staff at Mohawk and McMaster assisted with the conversion of OUAC and OCAS program codes into the seven generalized CIP code categories for their respective institutions. However, for the students who applied to other Ontario postsecondary institutions, due to the lack of consistency, we could not convert these codes into the same seven categories with fidelity. Instead, HEQCO researchers recoded each program code that appeared in the postsecondary application data files into one of two broad CIP categories: STEM or BHASE. This method enabled some degree of comparability with our seven generalized STEM and BHASE program categories, but the binary classification of postsecondary application program codes led to significant data loss.