Early-Onset Disability, Education Investments, and Social Insurance

By Robert Millard*

Working draft, please do not circulate

November 14, 2023

Click Here for Latest Draft of Working Paper

Abstract

Individuals with an early-onset (before age 18) disability attain substantially less education than their non-disabled counterparts. This equates to an eighteen percentage point gap in post-secondary attainment between these populations in Canada. This gap is related to how an early-onset disability affects the cost and return to investing in education and the availability of additional income through social welfare policies. I build and estimate a structural life-cycle model of education investment and labour market choices to analyze the effect of social insurance policies on education investments for early-onset individuals. I focus on two social insurance policies in Canada: provincial social assistance (SA) and federal disability insurance (DI). Using linked Canadian survey and administrative tax data, I estimate the model and reproduce the education gap, life-cycle employment rates, and attachment to SA and DI. I find the effect of disability on the accumulation and returns to human capital account for two-thirds of the education gap. However, 18.6% of the gap is related to disincentives from social insurance policies, mainly from added benefits in SA available for beneficiaries with disabilities. Through counterfactual experiments, I find decreasing the value of SA poses an insurance-incentive trade-off for early-onset individuals. Instead, post-secondary grants for early-onset individuals increase their educational attainment, employment, and improves welfare. Moreover, this policy helps pay for itself through added tax revenues and reduced dependence on SA.

*Stony Brook University.
I would like to thank Audra Bowlus, Nirav Mehta, Christopher Robinson, and Todd Stinebrickner for their supervision of this project and helpful feedback and comments. Thank you to Rory McGee, Lance Lochner, Tim Conley, Sergio Ocampo Diaz, Al Slivinski, Elizabeth Caucutt, Salvador Navarro, Lisa Tarquinio Baxter Robinson, Siha Lee, Timothy Moore, Steven Stern, David Wiczer, Hugo Benitez Silva, Marco Pariguana, and Tom Triveri. I am also grateful for helpful comments from participants of the 2022 Canadian Economics Association Annual Meeting, 2022 Midwest econometrics group conference, 2023 Association for Education and Finance Policy conference, 2023 Society of Economics of the Household conference, 2023 International Association of Applied Economists, and seminar participants at Stony Brook University, Queens University, the University of Windsor, and the University of Western Ontario. This article makes use of data administered by the Maytree foundation, Copyright 2023 Maytree. All rights reserved.

1 Introduction

Social insurance (SI) provides vital resources for individuals experiencing barriers to their gainful employment. However, this insurance comes with the well-known tradeoff of disincentivizing work and human capital investments in the labour market. In some cases, the distortionary incentives from SI may affect behavior before labour market entry, notably education investments, but less is known about these dynamic disincentives (Flinn and Mullins, 2015; Blundell et al., 2016). The behavioral incentives from SI are exacerbated for the population with disabilities, as disability can substantially lower the returns to work, and disability transfer payments are typically larger.[1] Understanding the insurance-incentive tradeoffs for the disabled population is policy-relevant, as rising caseloads threaten the fiscal solvency of disability programs (Autor and Duggan, 2006; David, 2011; Liebman, 2015; Milligan and Schirle, 2019). Addressing the moral hazard related to dynamic disincentives is one way to alleviate financial liabilities by increasing the education of potential applicants and their subsequent returns to work.

This paper investigates the dynamic disincentives of SI on education choices for individuals with a disability before age 19, defined as early-onset. Early-onset disabilities are present during primary and secondary schooling, which are critical periods of skill development and investment in human capital. The consequences of disrupting human capital accumulation early in life extend into adulthood, as education is a crucial determinant of financial independence and labour market success.[2] A lower perceived return to work discourages education investments, even more so when generous disability policies increase the value of not working. Consequently, an education gap exists between the populations with and without disabilities.[3] I study an application to Canada, where there exists an 18pp gap in post-secondary completion between individuals with and without an early-onset disability.

To quantify the dynamic disincentive of disability policy, I build and estimate a structural model of post-secondary education investments, labour supply, and application to SI. The model accounts for a rich set of channels by which disability can discourage education to isolate the disincentive effects of labour market disability policy. Education choices are dynamic in nature and made based on the expected returns to employability and earnings. First, education is an investment into one's productivity, and the expected return

[1] Disability onset in working life has been found to reduce labour force attachment, earnings, and consumption, and increase reliance on government transfers (Burkhauser et al., 1993; Bound and Burkhauser, 1999; Haveman and Wolfe, 2000; Millard, 2023)

[2] Early-onset disabling conditions have been found to stunt earnings growth, lower labour force attachment, and increase dependence on transfer programs (Currie, 2009; Case and Paxson, 2010; Lundborg et al., 2014; Almond et al., 2018; Prinz et al., 2018).

[3] For instance, Case et al. (2005) find individuals with chronic conditions by age sixteen have a 16 percentage-point gap in post-secondary completion compared to individuals without chronic conditions in the UK. Loprest and Maag (2007) report gaps of 17% and 21% in college/post-college graduation rates between early-onset and non-disabled individuals in the US.

to schooling is increasing in one's human capital. A disability can interfere with productive skill development during childhood and adolescence, resulting in lower ability when post-secondary choices are made (Heckman, 2007; Currie and Almond, 2011). Additionally, a disability can interfere with skill accumulation during post-secondary and in the labour market after that (Cutler et al., 2006; Mori, 2016). Second, education is a costly investment, and the presence of a disability can exacerbate the financial and psychological costs of post-secondary.[4] Finally, disability alters the labour market environment in ways that lower expected returns to work, discouraging schooling investments for forward-looking individuals. A disability can lower the expected likelihood of finding and maintaining employment.[5] Moreover, SI will discourage education if it raises the expected outside option to work. This last channel is especially relevant for individuals with an early-onset disability, given their lower productivity and the greater generosity of SI programs available to them. Moreover, this group has the potential to rely on SI for longer durations, given the timing of their disability.

The model is representative of the Canadian labour market and SI policy environment. I focus on the two main income assistance programs available for individuals with disabilities: disability insurance (DI) and social assistance (SA).[6] DI delivers monthly transfers that are proportional to an individual's average pre-application earnings. Additionally, DI considers education in determining an applicant's eligibility, which can add to disincentives (Government of Canada, 2022). SA is means-tested welfare, which allocates additional resources for beneficiaries with disabilities (SA-D). The population of Canadians that are potentially eligible for these disability programs is sizeable, with a quarter of working-aged individuals reporting some degree of activity limitation. A quarter of this population reports their activity limitations to have onset before they are nineteen years old. Disability rates have been rising over the past few decades in Canada, as well as most developed countries, posing a significant financial cost to social infrastructure.[7]

I estimate the model using the Longitudinal and International Study of Adults (LISA), a panel survey of Canadian households that contains rich information on health, education, and other demographic characteristics. LISA is linked to a 36-year panel of administrative tax records containing disaggregated measures

[4]Financial costs may be higher due to disability-related expenses, such as the need to pay for accommodations. Psychological costs include stressors related to coursework and social aspects post-secondary, which may also be higher in the presence of a disability (Druckman et al., 2021).

[5]This may be due to employer beliefs, institutional features, search behaviour, or the need for workplace accommodations (Acemoglu and Angrist, 2001; Kitao, 2014; Ameri et al., 2018).

[6]This Canadian policy environment is structured similarly to other developed nations. For instance, the counterparts to these programs in the United States are Social Security Disability Insurance and Supplementary Security Income, respectively.

[7]The percentage of Canadians aged 15 and over with a disability rose from 12.4% in 2001 to 22.3% in 2017. This trend is likely to continue with an aging population as disability risk tends to increase with age. This increase may also be partially due to the broadening of the criterion for disability and changes in individual reporting behaviour. For more details on the economic position of Canadians with disabilities, see Cossette and Duclos (2002) and Morris et al. (2018).

of personal incomes, taxes, and transfers, allowing me to separate the behavioural incentives of SA and DI. Additionally, the administrative tax data alleviates concerns of measurement error and under-reporting that are often associated with self-reported survey measures of income (Gallipoli and Turner, 2009; Meyer et al., 2009). The merged survey and administrative information facilitate the creation of a rich dataset with detailed health measures, demographics, and incomes. Furthermore, LISA contains data on life-cycle outcomes for a relatively large subsample of individuals with early-onset disabilities, which is needed for conducting research into the lifetime consequences of disability for this group.[8]

The estimated model reproduces the gap in educational attainment, along with other differences in labour market outcomes between early-onset individuals and their not early disabled counterparts. I recover an employment gradient with respect to education that is comparable with the observed data, and I fit a positive average earnings difference by education level. Early-onset individuals also have a lower financial return to education and work experience, representing the disruption of skill accumulation in post-secondary schooling and the labour market. Moreover, the model reproduces life-cycle rates of DI receipt, and the estimated DI acceptance probabilities are consistent with the unconditional acceptance probabilities observed in the 2015 audit of the Canadian Federal DI program (Office of the Auditor General of Canada, 2015).

An estimated structural model facilitates analysis of the behavioural relationships between education investments, labour market conditions, and the policy environment. I use the model to decompose the post-secondary education gap into a set of underlying factors. To illustrate, I can distinguish the role of policy from other factors by shutting down disability policy, resolving the model, and simulating how behaviour changes. Moreover, I use the model to evaluate the effects of counterfactual reforms to the policy environment by simulating changes in individual behaviour, welfare, and net costs to the government. The counterfactual approach compares the effects of policy reforms to the current state of the world, which serves as the baseline.

The decomposition exercise finds that factors relating to human capital are the most significant contributor, accounting for 65.2% of the education gap. Half of this factor's contribution is from the impact of early-onset disability on experience-related skill accumulation. Approximately one-third is related to early-onset disability disrupting skill accumulation during post-secondary. The remainder relates to the impact of disability on skill accumulation in childhood and adolescence. Second, I find that dynamic disincentives from SI policy are the second largest contributor to the education gap, accounting for 18.5%. The role of SI

[8]Studying early health conditions has historically been difficult as it requires data on early life conditions and information on the outcomes of interest, often from adulthood. LISA provides me with information on Canadians over much of their life cycle and identifies early-onset disability with a retrospective survey question.

policy is almost entirely due to SA-D benefits, which account for 90% of this factor. This is due to SA-D benefits raising the relative value of low-earning states and individuals with low potential earnings substituting away from post-secondary. DI does not substantially contribute to the education gap. This is because the expected value of DI is increasing in average labour market earnings, and early-onset individuals have lower average earnings. Moreover, the expected value of DI is heavily discounted at the time individuals make their schooling decisions, as people tend to apply for DI at older ages. Psychic costs to schooling, disability risk, and the effect of disability on preferences each account for approximately 4% of the total education gap.

Finally, I use the estimated model to evaluate the effects of a set of counterfactual policy reforms on the education attainment and life-cycle behaviour of early-onset individuals. Through these exercises, I evaluate the implications of the insurance incentive tradeoff of SI on education outcomes, individual welfare, and net costs to the government budget. The decomposition exercise identifies areas that policy can target to reduce the education gap. To facilitate comparisons across the reforms, I fixed a policy objective of increasing the post-secondary attainment of early-onset individuals by one percentage point. First, a policy lowering the generosity of SA-D incentivizes individuals to self-insure their income with more education, resulting in substantial savings for the government. However, reducing this resource comes with substantial welfare costs to individuals. This reform mechanically takes away resources from individuals who end up relying on SA-D despite the reform and who have the largest marginal utility of consumption. This counterfactual scenario highlights the insurance-incentive tradeoff of SA-D with respect to education.

A second set of policy reforms considers incentivizing education with increased government expenditure. Most notably, I consider providing an annual transfer to subsidize the consumption of early-onset individuals during post-secondary schooling. This policy offsets the psychic cost of schooling, which may prevent high-ability individuals from enrolling. This policy is revenue-positive for the government, as increases in tax revenues and reduced dependence on SA offset the subsidy costs. Moreover, this policy reform improves welfare for early-onset individuals relative to the baseline.

I frame the contribution of this research in three broad areas. First, I contribute to a sizable literature on the relationship between early-life health, education investments, and labour market outcomes.[9] Health conditions at young ages can impede one's development in ways that persist for one's entire life. My contribution complements these studies by distinguishing and comparing the relative importance of mechanisms, such as human capital or SI policy, underlying the education gap and analyzing how this relates to adult

[9] For examples and surveys of related studies see Currie (2009), Case and Paxson (2010), Lundborg et al. (2014), Mori (2016), Almond et al. (2018), and Prinz et al. (2018).

inequalities.[10] I emphasize the role of labour market policy in incentivizing higher education investments.

Second, I contribute to a body of literature on insurance-incentive tradeoffs of disability policy by accounting for a broader set of behavioural responses to these programs. This paper fits among several studies that structurally model how disability policy drives labour market behaviour.[11] Much of this literature focuses on later onset disabilities, taking education level as given. However, it is crucial to account for the incentives SI has on early life decisions when considering the early-onset population because of the timing of their disability. To my knowledge, this is the first study to measure an insurance-incentive tradeoff of disability policy with respect to educational investments.[12] My results offer important insights into the design of DI and welfare programs when considering people affected by an early-onset disability.[13] I find that behavioural incentives matter for more than labour market decisions. Moreover, the moral hazard arising from these dynamic disincentives offers insight into the causes of application to disability programs.

Third, I contribute to the literature on the relationship between human capital investments, labour market conditions, and SI policies. Again, this paper aligns with studies linking education rates to the labour market environment.[14] The idea is that risks and public policies create incentives that distort behvaiour in the labour market. If these distortions are large enough, they may also affect pre-entry decisions. Education is arguably the most important human capital investment decision before labour market entry. If the labour market distortions created by SA or DI are large enough, they can significantly impact the returns to schooling for this group. My research also relates to the literature studying how individuals make their education decisions given future uncertainty.[15] My contribution is to evaluate the role of SI policy in partially insuring against uncertainty, affecting the expected value of self-insurance against future shocks through investing in post-secondary schooling.

The remainder of this paper is organized as follows: Section 2 outlines the details of the Canadian policy environment. Section 3 describes the features of the dataset. Section 4 provides some motivating descriptive evidence from these data. Section 5 details the empirical model. Section 6 discusses the strategy to estimate the model's parameters. Section 7 reviews the estimation results. Section 8 uses the model to decompose

[10] Mori (2016) conducts a similar structural analysis of how an early-onset disability affects education investments but does not account for labour market risks or the disincentives from SI.

[11] For instance, Gallipoli and Turner (2009), Bound et al. (2010), Kitao (2014), Low and Pistaferri (2015), Michaud and Wiczer (2018), Kostøl et al. (2019), and Kellogg (2021).

[12] Deshpande and Dizon-Ross (2023) conduct a similar study into dynamic disincentives of disability policy in the United States. They developed an experiment that provides parents of children with disabilities with information about the Social Security Income program, the US counterpart to the SA programs in Canada. Their focus was on how parental investments in their children were affected by this information. In contrast, I am interested in the post-secondary choices of individuals with early-onset disabilities.

[13] This last point is relevant for theoretical literature on the design of SI policy, such as Golosov and Tsyvinski (2006)

[14] For instance, Flinn and Mullins (2015), Blundell et al. (2016), and Bobba et al. (2021).

[15] For instance, Carneiro et al. (2003), Cunha et al. (2005), and Navarro and Zhou (2017).

the education gap and conduct counterfactual experiments. Finally, Section 9 concludes.

2 Disability Policy Environment in Canada

The Canadian SI environment is comprised of a set of programs at both the provincial and federal levels. For individuals affected by disability, programs offer assistance related to income insurance for earnings lost because of a disability, rehabilitation or reintegration into the workforce, and welfare for individuals unable to provide for themselves (Torjman and Makhoul, 2016). The programs differ in eligibility requirements, the screening of the population covered, the duration of aid provided, and the amount of aid provided. The disability programs operate relatively independently from one another rather than jointly administered or unified in delivering support, as in other countries. While this feature is convenient for separately analyzing disability policies, critics have argued this independence results in gaps in support for individuals with disabilities.

This paper focuses on the two main programs providing long-term income assistance and replacement for individuals affected by disability.[16] These are the Canadian Pension Plan Disability (CPP-D), the federal DI program, and provincial SA programs, which offer means-tested welfare payments.[17] The Canadian Pension Plan (CPP) is the federal retirement pension program that administers CPP-D. This section describes the main features of DI and SA in Canada.

2.1 Canadian Pension Plan Disability

DI in Canada delivers monthly financial transfers to applicants who are assessed and deemed eligible for the program. Eligibility requires applicants to be under the age of 65, not currently receiving Canadian Pension Plan (CPP) retirement benefits, have made a predetermined number of contributions to CPP, and are markedly restricted by a physical or mental disability. Individuals must complete and submit an application, be deemed to meet the eligibility requirements, and wait approximately 120 days for their application to be processed and approved before becoming a beneficiary of CPP-D.

First, eligibility depends on the characteristics of the disability and its impact on labour market performance. To receive CPP-D, an applicant must first show that their disability is both prolonged and severe. A

[16] I do not focus on other programs related to disability support, such as transportation or prescription supports, as these are considered a distinctly different policy area (Torjman and Makhoul, 2016). Furthermore, in its current state, the paper does not account for short-term insurance programs from employment insurance. I also do not model worker's compensation, which is only available to individuals injured at work and is not accessed by early-onset individuals in my data.

[17] In the following, I use CPP-D and DI interchangeably.

disability is prolonged if expected to be indefinite or likely to result in death.[18] The severity of the disability concerns the applicant's ability to engage in "substantially gainful activity" in the labour market. That is, how productive a disabled individual is in a job they could be expected to hold given their qualifications relative to others doing the same work but who do not have a disability. Program adjudicators make a subjective assessment of an applicant's scope for substantially gainful activity given their disabling condition and determinants of productivity. Adjudicators consider an individual's age, education, and work experience (Government of Canada, 2022).

The second eligibility requires that applicants have contributed to the CPP in four of the previous six years.[19] Contributions to CPP are compulsory for working Canadians aged 18 to 70. Contributions equal a percentage of a worker's bounded employment earnings. In 2019, contributions equaled 4.95% of a worker's employment income up to $55,900 in that year (Government of Canada, 2023). The contributions to CPP determine the monetary value, or generosity, of the CPP-D payments. The contributory period begins at age 18 and ends at age 65 or the year of death. It excludes years in which the applicant was receiving CPP-D benefits.

The monthly generosity of CPP-D is a function of an earnings index summarizing the average monthly earnings in the applicant's contributory period. In the calculation of the earnings index, applicants can drop certain months from their contributory period, which would reduce their final amount of CPP benefits.[20] CPP-D payments are the sum of two components. The first component is equal to 75% of the applicant's potential CPP retirement benefits at the date of application. Potential CPP retirement benefits are equal to 25% of an earnings index that summarizes an applicant's bounded average earnings over their contributory period. The minimum bound to their earnings has been $3,500 per year since 1996, and the maximum, which was $53,600 in 2015, is updated yearly based on a measure of average wages. The second component is a deterministic flat-rate benefit indexed by the CPI each year.[21]

[18] CPP-D is a program for long-term disabilities and not designed to insure against short-term disability spells.

[19] Three of the previous six years if the applicant has contributed to the CPP for twenty-five years or more.

[20] First, each applicant is eligible to drop contributory months in which their children were under the age of 7. Second, applicants can drop a remaining percentage of their remaining contributory months with the lowest earnings. The percentage of months eligible to drop after the months when the applicant's children were under the age of 7 have been removed was 15% prior to 2012, 16% for 2012 and 2013, and 17% thereafter.

[21] In 2018, the average CPP-D benefit received was just under $1000 per month, half of which was the deterministic flat rate component (Employment and Social Development Canada, 2018).

2.2 Provincial Social Assistance in Canada

The main source of welfare transfers in the Canadian social safety net is provincially administered SA Programs. SA offers last-resort financial assistance to individuals with barriers to sustained employment and who have insufficient or volatile sources of income. As SA is social insurance of last resort, it is available only for individuals who have exhausted all other means of assistance. This implies that SA beneficiaries must have sought income support from other sources, including DI. An important difference between SA and DI is that SA programs do not have a work requirement, widening the population these programs cover.

SA is separately administered in each province. As such, the SA programs vary in eligibility criteria and the generosity of their transfer by province. However, all SA programs share a similar overall structure (Employment and Social Development Canada, 2016).[22] Applicants to SA must be assessed to be in need of financial aid, and the value of aid provided depends on the magnitude of this assessed need. The eligibility and generosity of aid are based on a means test of the applicant's assets, earning capacity, and demographic characteristics, such as health status.

The means test calculates the net difference between an applicant's "assessed needs" and their financial assets to determine eligibility. An applicant is eligible for SA if their assessed needs exceed the sum of their income and assets below an upper threshold. First, an applicant's "needs" may include variables like living expenses, family size and composition, and disability.[23] Assessed income combines all earnings from market activities, such as paid employment or self-employment, with transfers from other government programs, such as DI. Individuals may receive SA while earning from other sources, but this may reduce benefits according to the program's replacement rate. SA may be revoked if sufficient effort is not taken on the part of the beneficiary to receive other sources of income support.

Recipients to SA typically receive monthly financial transfers equalling a basic assistance amount and, in some cases, a special assistance amount. The basic assistance amount covers the basic costs of living, such as food, shelter, and clothes. The cost of living varies with an individual's demographic characteristics, notably their disability status. A disability may require additional expenses due to greater costs of living and barriers to employment. For these reasons, all SA programs have additional resources available for individuals affected by a disability, SA-D.

[22]SA programs have been criticized for lacking available information about their provisions, eligibility, and administration details. This lack of transparency creates difficulties for potential applicants and analysts, as discussed in (Kneebone and White, 2015; Béland and Daigneault, 2015; Torjman and Makhoul, 2016).

[23]On the other side of the mean-test, the applicant's financial assets include liquid assets, such as cash or convertible assets, and fixed assets, such as property. Exempt assets include those used for employment or transport, such as tools or automobiles, and assets related to savings plans used for education purposes, such as registered education savings plans. The combined fixed and liquid assets must not exceed a predetermined threshold, which varies by provincial jurisdiction. Additional details on SA programs can be found in Employment and Social Development Canada (2016) or Hillel et al. (2020).

3 Data: The Longitudinal and International Study of Adults

LISA is a panel survey of over 11,000 Canadian households over four biennial survey waves, starting in 2012. LISA covers a broad range of topics, including health, education, labour, social participation, and income. These data are supplemented with administrative records. Most notable are T1 family files (T1FF), which contain rich disaggregated measures of personal income from individual income tax filings. Many questions in LISA are retrospective, allowing me to build a comprehensive history of incomes and transfers between 1989 and 2017 for observations in my sample. Importantly, this sample includes a relatively large number of early-onset individuals with linked T1FF tax data. Combining data on early-life health with data on life-cycle incomes and labour outcomes makes LISA especially well-suited to this research agenda.

I use the survey waves of LISA to obtain the majority of demographic information for individuals. Each survey wave contains information about education level, labour market status, change in labour market status since the previous wave, job search activities, reasons for job loss, and details about limitations to daily activities, which are used to derive disability status. I derive education level by the self-reported highest completed certificate. I flag an individual as low education if their highest completed certificate is equivalent to high school or less. I flag individuals as post-secondary if they have completed any post-secondary certificate.[24] Individuals with post-secondary have made an additional human capital investment beyond compulsory education.

Measuring Disability

The 2014, 2016, and 2018 waves of LISA include measures of limitations to activities of daily living (LADL) and other characteristics of health, which are used to derive an individual's history of disability status.[25] The set of LADL's are derived from a short version of "the disability screening questions" (DSQ) developed by Statistics Canada for identifying individuals with disabilities in general population surveys (Grondin 2016). This model distinguishes five main areas of activity limitation: Seeing, Hearing, Physical, Cognitive, and Mental Health. Physical combines limitations to mobility, flexibility, dexterity, and pain. Cognitive disabilities combine developmental disabilities, limitations to learning, such as dyslexia or hyperactivity, and limitations to memory and concentration.[26] Mental health conditions encompass many emotional, psychological, and mental health conditions, including anxiety, depression, bipolar disorder, substance abuse,

[24] Post-secondary includes 2-year and 4-year college degrees, any university degrees, as well as vocational degrees.

[25] The 2012 wave comprises only a small set of questions about disability and excludes information on the age of disability onset.

[26] It is important to note that developmental disabilities such as Down syndrome, Autism spectrum disorder, Asperger syndrome, or brain damage due to lack of oxygen at birth typically manifest early in life rather than as late-onset disabilities.

and anorexia.[27]

The activity limitations are self-reported in LISA. For each type of activity limitation, respondents are asked a flow of categorical questions about the magnitude of difficulty and frequency of limitation for each LADL.[28] The short version of the DSQ flags disability based solely on the reported frequency of limitation. A respondent is flagged for a type of disability if reporting their condition to limit their activities "sometimes," "often," or "always."[29] The age of disability onset is derived from a self-reported retrospective question,"at what age did you first start having difficulty or activity limitation?" Due to the retrospective nature of this question and the panel structure of the survey waves, there are instances where an observation reported different ages of onset. To address this, I use the minimum reported age of onset as the truth.

Much research in health economics has focused on the validity of self-reported measures of one's health. One concern relates to the inherent subjectivity of how one assesses one's own health. For example, two otherwise identical individuals may differ in the reported severity of their disability. Additionally, critics of self-reported health measures argue that individuals may exaggerate the existence or severity of their health condition to justify poor economic outcomes or attachment to government programs, a phenomenon referred to as justification bias. The evidence on the endogeneity of self-reported health measures and the extent of measurement error are mixed (Black et al., 2017). Although, it is important to note that recent articles tend to find evidence for state-dependent reporting.[30]

My disability measure is derived from a respondent reporting any positive limitations to a specified activity and abstracts from the degree of impairment. This approach mitigates concerns related to subjectivity in the scale of impairment from a self-reported activity limitation, as I do not distinguish conditions along the severity margin. Moreover, much of the evidence on justification bias is based on broad questions about one's health or disability, such as "do you have a medical or physiological condition that impairs the type or amount of work you can do." The questions about activity limitations in this survey are linked to specific tasks, such as walking on a flat surface for fifteen minutes, grasping a small object like scissors, or experiencing ongoing memory problems or periods of confusion. Additionally, the presence of some activity limitations is

[27] More details on the survey questions can be found in Section A.2. of the Appendix.

[28] Some cognitive conditions, such as developmental disability or learning conditions, are initially flagged based on diagnosis from medical professionals instead of the level of difficulty. Refer to Section 1 in the Appendix for details.

[29] I flag disability based on the frequency of limitation alone, as there are inconsistencies in questions about the magnitude of difficulty across the survey waves.

[30] It has been found that self-reported disability is close to exogenous, may actually under-represent the extent disabled population, and may even underestimate the true impact of disability on relevant labour market outcomes (Stern, 1989; Bound and Burkhauser, 1999; Burkhauser et al., 2002). Others have found evidence of justification bias related to labour market states inflating the prevalence of health conditions (Benítez-Silva et al., 2004; Baker et al., 2004; Black et al., 2017). Moreover, alternate approaches to identify individuals with disabilities, for instance, by using disability insurance beneficiaries to define the population with a disability, have been found to under-represent the population of individuals who are limited enough in the labour market to be classified as "disabled" (Bound, 1989)

elicited based on whether the respondent has been diagnosed with a specific condition, such as a learning or developmental disorder, by a healthcare professional.[31] Last, mental health is identified using specific examples of diagnoses, such as anxiety, depression, bipolar disorder, or anorexia. These approaches narrow the scope of justification bias to be anchored to the activities in question, base the existence of a limiting condition on the diagnosis of a medical professional, or frame limitations related to mental health with specific examples of diagnoses. I follow much of the related literature and take the responses to questions on limitations to daily activities as given. However, I acknowledge the empirical concerns that are inherent to any self-reported measures of health.

T1 Family Files

The income histories come from the T1FF records. The T1FF spans from 1982 to 2017 and is linked to each respondent in the main survey waves of LISA. These data contain details on an individual's demographic characteristics relevant to their tax filings, such as age, marital status, province of residence, and the number of children. For this analysis, I focus on measures of employment income and government transfers from CPP-D and SA payments. I focus on 1989-2017 calendar years due to limitations on the income measures in the years prior.[32] A notable advantage of these tax records is that they are less likely to suffer from the measurement and coverage issues often associated with survey data. For instance, Meyer et al. (2009) show that survey measures of public transfers often suffer from respondents under-reporting, which can lead to overestimation of total income declines following the onset of disability.

3.1 Supplementary Data: Education Expenditure, Social Assistance and Taxes

Data on the maximum amount of SA are drawn from the annual report series, "Welfare in Canada," produced by the Maytree Foundation (Maytree, 2018).[33] These reports calculate the maximum amount from SA a household may receive in their respective province and calendar year, plus the maximum amount of provincial and federal tax credits available to individuals. These thresholds are calculated for four distinct household types. In this study, I use the maximum amount of SA for a single employable adult and for a single adult with a disability.[34]

The parameters of the Canadian income tax system are derived using data from the Canadian Tax and

[31] This type of question has been used to assess the validity of self-reported health measures in Baker et al. (2004)

[32] For instance, CPP-D is not separated from CPP, and SA is not separated from other nontaxable income in earlier years.

[33] These annual reports were formerly conducted by the National Council of Welfare until 2009.

[34] The TIFFs include an individual's reported income from SA and their taxes paid in a given year. I do not use this measure of SA in estimating the model as it depends on variables that I am not able to observe, most notably savings and assets.

Transfer Simulator (Milligan, 2016). This rich resource reports income tax brackets and respective marginal tax rates for the federal and provincial levels since 2016. Combining the federal and provincial rates gives a potentially distinct tax regime for each province in each calendar year covered by my study.

Finally, data for expenditures during education come from the Canadian Tuition and Living Accommodation Costs survey. This survey collects data for full-time students at publicly funded Canadian degree-granting institutions. I set consumption during post-secondary schooling to the average price-adjusted tuition and ancillary fees from 1993 to 2018, weighted by density over these years in the data.

Sample Selection

My sample of interest consists of males with and without an early-onset disability. I restrict the age in the sample to be between the ages 18 and 65 during the calendar years 1989 to 2017. LISA excludes individuals with disabilities who are institutional residents.[35] This population has very severe disabilities that eliminate any hope for their participation in the labor market. This project is interested in individuals with disabilities on the margin of choosing post-secondary. Finally, I exclude individuals living in the Canadian Territories.

4 Motivating Statistics

Table 1: Likelihood of Post-Secondary Attainment by Early Disability Status

	Data
Early-Onset	0.460
	(0.037)
Not Early Disabled	0.640
	(0.012)

Estimates are from LISA and T1FF over 1989-2016 and weighted to represent the of Canada population in 2012. Post-secondary education equals one if the individuals has completed any post-secondary, which includes college certificates, university degrees below a bachelors, a bachelors degree, and degrees above a bachelors. Individuals who complete high school or drop out are grouped into the low schooling category. Standard errors are reported in parenthesis below.

I illustrate the correlations between disability, education, and labour market outcomes in the raw data to motivate the relationship between education and SI. Table 1 shows that the likelihood of completing post-secondary is eighteen percentage points lower for early-onset individuals. Less than half of individuals affected by an early-onset disability complete a post-secondary degree. Education is a costly investment into one's productivity, and the return to education is reflected in labor market earnings and employment. The

[35] Institutional residents are individuals in general hospitals, prisons, nursing homes, and special care facilities for individuals with disabilities.

presence of a disability can inhibit the use and accumulation of human capital, resulting in negative effects on earnings and employment.

Table 2: Employment and Earnings by Education Level and Early Disability Status.

	Not Early Disabled		Early-Onset	
	Low Education	Post-Secondary	Low Education	Post-Secondary
Over All Years in Labour Market				
Annual Earnings($)	32300	50900	26000	40400
	(21300)	(31600)	(19900)	(27400)
Employment Rate	0.740	0.846	0.508	0.753
First 3 years in Labour Market				
Annual Earnings ($)	15100	20700	12900	18200
	(10800)	(14300)	(10000)	(13300)
Employment Rate	0.810	0.862	0.579	0.815

Estimates are from LISA and T1FF over 1989-2016 and weighted to represent the of Canada population in 2012. Post-secondary education equals one if the individuals has completed any post-secondary, which includes college certificates, university degrees below a bachelors, a bachelors degree, and degrees above a bachelors. Individuals who complete high school or drop out are low-education category. Standard deviations are reported in parenthesis below.

Outcome gaps by education level reflect how post-secondary schooling augments labor market productivity for those choosing to go while acknowledging that selection into higher education depends on many other determinants of earnings, such as ability and motivation.[36] That said, drawing comparisons in outcomes by early disability and schooling level is informative of how early onset disability affects one's labor market position, given their education level. Table 2 contrasts the difference in average lifetime earnings and employment by early disability status and education level. First, early-onset individuals with low education who work earn a fifth less than their not early disabled counterparts, placing them at greater risk of SI application. The low returns to working for early onset individuals are reflected in their lifetime employment rate, which is 23pp lower than not early disabled individuals with low education. The third and fourth rows of Table 2 compute average earnings and employment rates in the first three years after labor market entry. The magnitude of the differences by early disability status for low education is smaller, suggesting

[36]Note that these unobservable predictors of education may also be influenced by a disability. For example, it may require additional motivation to overcome the cost of a disability and pursue a bachelor's degree.

early-onset disability adversely affects the accumulation of skills in the labor market.

Table 3: Labour Market Uncertainty by Education Level and Early Disability Status

	Not Early Disabled		Early-Onset	
Labour Market Friction	Low Education	Post-Secondary	Low Education	Post-Secondary
Job Arrival Rate	0.730	0.759	0.650	0.713
	(0.000)	(0.004)	(0.000)	(0.000)
Job Destruction Rate	0.031	0.021	0.065	0.033
	(0.016)	(0.025)	(0.006)	(0.003)
Risk of Retirement	0.061	0.065	0.053	0.070
	(0.050)	(0.039)	(0.009)	(0.007)

Estimates are from LISA and T1FF over 1989-2016 and weighted to represent the of Canada population in 2012. Post-secondary education equals one if the individuals has completed any post-secondary, which includes college certificates, university degrees below a bachelors, a bachelors degree, and degrees above a bachelors. Individuals who complete high school or drop out are low-education category. Standard errors are reported in parenthesis below.

The earnings gap by early-disability status is similar in magnitude for those with post-secondary education, relative to those with low education. The smaller fraction of early-onset individuals choosing post-secondary education earn one-fifth less than their non-disabled counterparts, suggesting they have lower average ability, less of a financial return to their education, or both. Average employment for early-onset individuals with post-secondary is much closer to their non-disabled counterparts, as they have higher returns to work.

Observed employment rates and earnings reflect both the labour supply of individuals and the labour demand of employers.[37] Table 3 compares the rate of job offers for individuals who are searching and the rate of job destruction due to firing or layoff. Early-onset individuals receive fewer job offers, conditional on searching, regardless of their education level. Furthermore, early-onset individuals are displaced from work at a greater rate. Jobs arrive at a greater rate and are destroyed at a lower rate with a post-secondary degree. These rates are consistent with a larger set of more permanent/ stable jobs being available for those with post-secondary schooling. The last row in Table 2 reports the rate at which people retire between the ages of 60 and 65. The rate of retirement is larger for individuals with post-secondary, which is consistent with this group having accumulated enough wealth through higher earnings to retire early (higher CPP). Early-onset

[37]Information on employers is limited in the T1FF files. However, the survey waves of LISA identify a monthly transition rate into employment for individuals actively searching. Additionally, I calculate a job loss rate based on individuals who were observed to work in the previous survey and reported to have been fired or laid off.

individuals with high school or less are least likely to retire early, which is consistent with this group having relatively fewer available resources in retirement. This could result from a higher lifetime dependence on government transfers for these individuals.

In sum, early-onset individuals expect to have less of a return to lifetime earnings from going post-secondary relative to their non-disabled counterparts. This population faces greater search frictions that increase their likelihood of being out of work. Furthermore, this group enters working life with a disability, granting them more generous SI transfers. DI and SA are most relevant for early-onset individuals with low education, as this group is most exposed to low and volatile earnings and a higher risk of unemployment. This idea is reflected in Table 3, which shows the likelihood of receiving transfers and the average transfer amount from DI and SA by early-disability status and education level.

In rows 1 and 2 of Table 4, a stark difference exists in SA rate by early disability status. Early-onset individuals are most likely to become a recipient of SA early in their life and receive larger benefits from SA on average. The percentage of people ever becoming a beneficiary of SA is more than twice as large for the early-onset group across all education levels. Over thirty percent of the low-education group is dependent on the program at some point in their life. Additionally, the difference in benefits received between early-onset and not early disabled, shown in rows 3 and 4, is decreasing in education, with the low education group receiving \$2,000 more per year when they are early-onset and approximately \$800 more per year on average for not early disabled.

Rows 5 and 6 of Table 4 report that the likelihood of receiving CPP-D is low, culminating to about 4% of early-onset individuals becoming a beneficiary of the program. It is important to note that this number represents the individuals who applied and were accepted to CPP-D. It may be the case that many more people apply but are not accepted. In 2014-2015, 43% of total applications were accepted to CPP-D (Office of the Auditor General of Canada, 2015). The size of DI benefits received increases with age as lifetime earnings grow.

This evidence suggests a much larger portion of the early-onset population is likely to require some form of government assistance in their life, especially those with lower levels of education. Lower levels of education are associated with greater labour market risk, substantially lower earnings capacity, and greater volatility in labour force attachment. The combination of these findings is consistent with early-onset individuals facing different incentives from labour market risks and policies when making their educational investments. To measure and distinguish between the size of the effect of policies and risk on educational investments, I next build and estimate a life-cycle model of educational investments and behaviour in the labour market.

Table 4: Average Rate and Transfer Amount From Social Assistance (SA) and Disability Insurance (DI) by Education Level and Early Disability Status

	Not Early Disabled		Early-Onset	
	Low Education	Post-Secondary	Low Education	Post-Secondary
SA Rate				
Age < 45	0.0773	0.0252	0.3702	0.0772
	(0.003)	(0.001)	(0.014)	(0.006)
Age ≥ 45	0.0785	0.0262	0.2963	0.1309
	(0.003)	(0.001)	(0.019)	(0.013)
Average Transfer from SA				
Age < 45	6100	5600	8200	6800
	(100)	(200)	(200)	(300)
Age ≥ 45	7200	6600	8700	6100
	(100)	(200)	(300)	(300)
All Labour Market Years				
DI Rate	0.0238	0.0085	0.0396	0.0407
	(0.001)	(0.000)	(0.005)	(0.004)
Average Transfer from DI	9100	9300	7600	7800
	(100)	(100)	(200)	(200)

Estimates are from LISA and T1FF over 1989-2016 and weighted to represent the of Canada population in 2012. Post-secondary education equals one if the individuals has completed any post-secondary, which includes college certificates, university degrees below a bachelors, a bachelors degree, and degrees above a bachelors. Individuals who complete high school or drop out are low-education category. Standard errors are reported in parenthesis below.

5 Empirical Life-Cycle Model

The descriptive evidence motivates a behavioural relationship between SA, DI, and education. To analyze the factors driving this relationship, I develop a life-cycle model of education investments and labour market decisions that is representative of the Canadian environment. Specifying a structural model allows me to analyze the mechanisms underlying education choices for early-onset. The model formalizes a link between the net returns to education, the labour market, and SI policy, which I use to compare the relative importance of factors underlying the observed education gap between early-onset and their not early disabled counterparts. Furthermore, the model allows me to analyze the effects of reforms to the policy environment. This provides a broader understanding of behavioural responses to policy and is an important step in considering the normative consequences of reforms on social and individual welfare.

5.1 Model Preliminaries and Initial Conditions

Time is discrete, and each period represents a year. Individuals enter the model at t=0, corresponding to 18 years of age, and choose to go to post-secondary (s=1) or enter the labour market out of high school (s=0). This choice depends on a set of endowments that affect the expected return to each education level. Initial endowments include disability status, $d_0 \in \{0, 1\}$, which identifies the early-onset group. Individuals with $d_0 = 1$ have an early-onset disability and $d_0 = 0$ otherwise. I use a unitary notion of disability in this paper. Disability matters by lowering returns to education and work and giving eligibility to SI, which raises the value of not working. The direction of these effects is true regardless of the type of disability. Given d_0, individuals receive an ability endowment, $a_i \sim N(\bar{a}^{d_0}, \sigma^2_{a^{d_0}})$ and and a preference shock to schooling, $\psi \sim N(\bar{\psi}^{d_0}, \sigma^2_{\psi})$. The ability endowment interacts with education in determining earnings at labour market entry.

The initial unobserved heterogeneity is restricted to impact education choices through two channels. The ability endowment represents all unobserved factors influencing skill development before age 18. Psychic costs capture many factors affecting the monetary cost and preference of post-secondary schooling but do not affect productive ability. For instance, average differences in tutoring investments during adolescence by d_0 will affect the ability distribution. Whereas variation in parental resources by d_0, which affect the financial cost of post-secondary education, are reflected in the distribution of the psychic cost of schooling. Other factors, like parental education, influence both ability development in adolescence and the preferences for post-secondary education and will be reflected in a_i and ψ_i.

Education level is chosen to maximize an individual's expected discounted lifetime utility. The expectation is a function of the labour market environment, including the policy environment and a set of risks that may differ by education and early disability status. Those choosing post-secondary enter the labour market at 22 years old. Time in the labour market lasts until, at most, age 65, upon which everyone faces ten mandatory periods of retirement and then dies.[38] Additionally, individuals face an exogenous risk of early retirement when they turn 60, which depends on d_0 and s. The lifespan of 75 years of age (T=57) is fixed for all individuals, and I assume there is no bequest motive. The total lifetime can be split into time in school, T^S, time in the labour market, T^L, and time in retirement, T^R.

[38]The model assumes there is no mortality risk before the terminal period. Empirically, mortality rates are trivially low before age 60 and, therefore, heavily discounted for the schooling decision. Moreover, differences in mortality by early disability status and education level only appear after age 60 and are small in magnitude.

5.2 Labour Market Environment

I first outline the structure of an individual's decision problem in the labour market. Individual i's observed choices at age $t \in T^L$ depend on their initial endowments and time-varying state variables, which evolve according to labour market risks given the sequence of choices before t. Individuals choose to work and earn employment income, not work and receive SA, or to apply for DI, and hence can find themselves in one of three labour market states: working, not working on SA, or not working and on DI. Individuals make their decisions subject to uncertainty in their future disability status, finding and maintaining a job, and productivity.

Disability Risk

Disability status, $d_{it} \in \{0, 1\}$, evolves according to a first-order Markov process, where $d_{it} = 1$ when disabled in period t and $d_{it} = 0$ otherwise. Disability risk is assumed to be exogenous to an individual's labour market choices.[39] Disability transition probabilities vary with age and initial disability status. The risk of disability onset increases with age, and the likelihood of recovery decreases with age. The disability transition probabilities vary by d_0, as early-onset disabilities represent a potentially different set of conditions that are allowed to evolve differently over the life-cycle. The transition probability for disability status is defined as

$$\gamma_{k,l}^{d_0,t} = Pr(d_t = k | d_{t-1} = l, t, d_0), \; k, l \in \{0, 1\}. \tag{1}$$

Search Frictions

While not working, individuals may decide to enter employment if they receive an offer with probability, $\lambda^{d_0,s}$. While employed, an individual who does not choose to leave employment is exogenously displaced out of employment with probability, $\delta^{d_0,s}$. Employed individuals may also choose to quit their jobs endogenously. These frictions are allowed to vary by early-onset disability status to account for differences in search behaviour, institutional features, employer beliefs, and other barriers to working for this group.[40]

[39]Modelling disability risk this way is standard in the related literature (Low and Pistaferri, 2015; Michaud and Wiczer, 2018; Kostøl et al., 2019; Kellogg, 2021)

[40]Articles motivating the inclusion of these risks include Acemoglu and Angrist (2001), Dixon et al. (2003), Kitao (2014), and Morris et al. (2018).

Annual Earnings

An individual's potential earnings are determined by a combination of potential work experience, PE_{it}, current and initial disability status, d_{it} and d_{i0}, time-varying idiosyncratic shocks to their productivity, $\epsilon_{it}^{d_0,s}$, and unobserved fixed heterogeneity, $v^{d_0}(s_i, a_i)$. Potential earnings in an arbitrary period, t, are

$$\ln W_{it} = \mu_1^{d_0,s} PE_{it} + \mu_2^{d_0,s}(PE_{it}/100)^2 + \phi d_{it} + v^{d_0}(s_i, a_i) + \epsilon_{it}^{d_0,s} \tag{2}$$

$$\begin{aligned} &\text{where } \epsilon_{it}^{d_0,s} = \epsilon_{it-1}^{d_0,s} + \xi_{it}^{d_0,s}, \\ &\quad \xi_{it}^{d_0,s} \sim N(0, \sigma^2_{\xi^{d_0,s}}) \text{ for } t > 0, \\ &\text{and } \xi_{i0} \sim N(\bar{\xi_0}, \sigma^2_{\xi_0}). \end{aligned} \tag{3}$$

The parameters governing potential earnings depend on initial disability status and education level. The second-order polynomial of experience provides curvature to the life path of potential earnings. The specificity of $\mu_1^{d_0,s}$ and $\mu_2^{d_0,s}$ to initial disability status lets d_0 and s to affect the evolution of earnings over the life-cycle. The return to potential experience also varies by education level, representing heterogeneity in the rate of productive skills accumulation on the job.[41] The direct effect of a disability on productive human capital is captured by ϕ. This encompasses a disability-induced loss of work-relevant skills, negatively affecting earnings.

Permanent productivity shocks, $\epsilon_{it}^{d_0,s}$, follow a random walk with identically and independently distributed innovations, $\xi_{it}^{d_0,s}$. These shocks reflect that volatility in earnings may differ by initial disability status and education level. These can be interpreted, for example, as shocks to the value and price of individual skills or as disability bias technological change, which impacts the set of feasibly productive jobs. I assume $\xi_{it}^{d_0,s}$ is normally distributed with mean zero and variance $\sigma^2_{\xi^{d_0,s}}$.

An early-onset disability also impacts the development of productive skills during school. Unobserved fixed heterogeneity, $v^{d_0}(s_i, a_i)$, can be interpreted as an individual's human capital upon entry to the labour market given their education. To capture differences in the return to education by early-onset disability, I make the following parametric assumption:

$$v^{d_0}(s_i, a_i) = \begin{cases} h_0^{d_0} a_i + \xi_{i0} \text{ if } s_i = 0 \\ h_1^{d_0} a_i + \xi_{i0} \text{ if } s_i = 1. \end{cases} \tag{4}$$

[41] For instance, Cutler et al. (2006) study heterogeneity across education levels in one's ability to cope with a disability and its effect on the evolution of their life-cycle earnings.

The parameter $h_s^{d_0}$, $s \in \{0,1\}$ scales an individual's ability endowment differently depending on their initial disability status and chosen schooling level.[42] Initial earnings also depend on an initial human capital shock, ϵ_{i0}, representing productive human capital that is unrelated to education.

The Earnings Index

The earnings index, a summary measure of an individual's earning history in T^L, is used to determine the generosity of DI and retirement transfers.[43] The earnings index, e_t, is assumed to update each period given the previous period's earnings index, e_{it-1}, the individual's labour earnings in the current period, W_{it}, and age, t, according to

$$e_{it} = f(e_{it-1}, W_{it}, t) = \begin{cases} \frac{(t-1)e_{it-1}}{t} & \text{if } W_{it} < \underline{W} \\ \frac{(t-1)e_{it-1}+W_{it}}{t} & \text{if } W_{it} \in [\underline{W},\ \bar{W}) \\ \frac{(t-1)e_{it-1}+\bar{W}}{t} & \text{if } W_{it} \geq \bar{W}, \end{cases} \tag{5}$$

where $e_{it} = 0$ for $t \in T^S$. The parameters $\underline{W}$ and $\bar{W}$ are the lower and upper bounds, respectively, on average earnings in period t. These are set to $\bar{W} = \bar{W} = \$40,000$ and $\underline{W} = \underline{W} = \$3,500$, which reflect the actual value of the upper and lower bound used in these policies. If not employed, the earnings index updates according to $W_{it} = 0$, inducing a cost to non-participation in the labour market.

Retirement

Individuals face a retirement risk, $\rho^{d_0,s}$, starting at age 60, which differs by initial disability status and education level.[44] Retirement income comes from pension benefits and old age security. Retirement benefits equal $0.25 * e_{it}$, which approximates the formula used in the CPP. Old age security is fixed at \$5,500, which approximates the average amount received from the Old Age Security Pension (OASP) program in Canada.[45] OASP helps supplement income for retirees with no CPP income.

Individuals may be exogenously shocked into retirement starting at age 60. Retirement risk, $\pi_{ret}^{d_0,s}$, depends on education level and early disability status. If retiring early, an individual's retirement income is penalized 7.2% for each year they are retired before age 65, up to a maximum of 36% for those who retire

[42] This specification is similar to Flinn and Mullins (2015) and has the feature that human capital production technology is supermodular in ability.

[43] This index is similar to the averaged indexed monthly earnings measure that is a determinant of the Supplementary Security Income and Social Security Disability Insurance programs in the US.

[44] Education and early-onset disability may affect determinants of early retirement. For instance, d_0 affects the likelihood of having a disability at older ages.

[45] This value is consistent with the average OASP income received observed in the T1FF.

at age 60. The penalty lasts for the duration of their retirement, and individuals do not receive \$5,500 from OASP until age 65.[46]

Disability Insurance

The DI program in my model is intended to approximate CPP-D. DI provides partial insurance to individuals who are under the age of 65, are restricted in their ability to engage in any substantial gainful activity due to their disability, and who meet the program's contribution requirements. I make the following simplifying assumptions to the DI program for computation tractability. First, eligibility for DI relies on the interaction between an individual's disability status and their productivity in the labour market, which defines what is deemed substantially gainful activity and is imperfectly observed. Hence, DI is awarded to applicants with error, and DI acceptance is modeled probabilistically. DI administrators use an applicant's observable characteristics, such as their education, to gather information about whether the applicant is unable to engage in any substantially gainful activity. Hence, the acceptance probabilities vary with s.

To approximate the contribution requirement of CPP-D, I assume that individuals must have worked at least once to be eligible for DI. This requirement is captured by the binary variable $elidg_{it}$, which equals one if the contribution requirement is met and equals zero otherwise.[47] I assume DI is only available for those with $d_{it} = 1$.[48] Hence, conditional on having applied to DI in the previous period, $m_{it-1} = 1$, the probability of acceptance is

$$\mathrm{PR}(\mathbb{1}_{it}^{DI} = 1 | elidg_{it} = 1, d_{it} = 1, s_i) = \pi^s. \tag{6}$$

An individual's CPP retirement benefits are approximated as 25% of their earnings index, e_{it}. DI benefits are equal to 75% of their CPP retirement benefits plus the flat rate component.[49] The yearly flat rate component of DI of \$4,365 is set by policymakers and known to agents in the model.[50] Hence, DI generosity is given by

$$DI_t(e_{it}, b) = 0.1875\ e_{it} + 4,365. \tag{7}$$

[46] If their early retirement income falls below the amount of SA they are eligible for, their income is topped up with SA.

[47] This assumption has bite if individuals seeking DI with no work history work for one period in order to meet the contribution requirement. However, this behaviour does not occur in estimating and simulating the model.

[48] As disability is measured based on limitations to daily activities, this assumption may miss some individuals with a health condition that automatically grants them access to DI. The sample of individuals who never report a disability but end up on DI in the data is trivial in the data and only occurs at the very end of the life-cycle.

[49] I model DI generosity in a similar manner as Gallipoli and Turner (2009) and Milligan and Schirle (2019).

[50] The real value of this amount has fluctuated between \$3,900 - \$4,500 over the calendar years spanned by the T1FF. The flat rate component reflects a weighted average of this value over the years covered by my sample.

I also set the maximum amount to \$40,000, and the yearly basic exemption is \$3,500, which is representative of the true program.[51]

Individuals face a utility cost of applying to DI, $C_{App}^{d_0,s}$. The application process can be lengthy and requires the applicant to compile a set of eligibility resources. The psychic cost of this process may differ by schooling, which can help one with the skills to gather this set of resources. Alternatively, education level may be correlated with an individual's preference for self-sufficiency in the labour market. I allow this disability cost to differ by d_0, as early-onset individuals may be more familiar with the disability social safety net or have different preferences for self-sufficiency in working life.

Lastly, individuals face a risk of reassessment of eligibility, ρ, for DI. If benefits are reassessed, individuals' benefits are terminated, and they will need to re-apply to become a beneficiary again.

Social Assistance

SA benefits are means-tested anti-poverty programs. The maximum amount of benefits from SA programs differs by disability status. These programs are intended to ensure that the income of individuals does not fall below a specified threshold. In the model, I approximate provincial SA programs and the determination of SA benefits. I assume the lower bound on consumption, $\bar{c}(d_{it})$, depends on disability status, representing the added SA-D resources for recipients affected by disability. I assume that there is 100% take-up of this program when not working or on DI. I define inc_{it} as an individual i's income from all other sources. Then, the formula for SA is

$$SA(inc_{it}, d_{it}) = \begin{cases} \bar{c}(d_{it}) - inc_{it}, & \text{if } inc_{it} < \bar{c}(d_{it}) \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

The allowance of the additional SA-D benefits is granted with probability, π^{SA}. I define $\mathbb{1}^{SA-D} = 1$ if $d = 1$ and approved for SA-D, so that

$$\bar{c}(d_{it}) = \begin{cases} 6,464 & \text{if } \mathbb{1}^{SA-D} = 0 \\ 9,786 & \text{if } \mathbb{1}^{SA-D} = 1. \end{cases}$$

SA and SA-D are calculated by the weighted average over all province year thresholds reported for single employable adults and single adults with a disability, respectively, in the Maytree Social Assistance Reports (Maytree, 2018). The weights are equal to the density over the years and provinces observed in my data.

[51] This assumption is reasonable as the flat rate component increases based on the Consumer Price Index. Hence, individuals expect the flat rate component to maintain the same amount of purchasing power across time.

Preferences

I assume a constant relative risk aversion (CRRA) utility function where consumption is non-separable from work and disability status.[52] The utility functions for working (W) and non-working (N) individuals are given by

$$U(c_{it}, L_{it}; d_{it}) = \begin{cases} u^W(c_{it}; d_{it}) = \frac{(c_{it} e^{\theta d_{it} + \eta})^{1-\kappa}}{1-\kappa} & \text{if } L_{it} = 1 \\ \\ u^N(c_{it}; d_{it}) = \frac{(c_{it} e^{\theta d_{it}})^{1-\kappa}}{1-\kappa} & \text{if } L_{it} = 0. \end{cases} \tag{9}$$

This specification implies that disability and work may affect the marginal utility of consumption. I assume θ and η are negative, which implies that workers or individuals with a disability require higher levels of consumption to have the same utility as non-working or non-disabled individuals.[53] These parameters capture the utility loss induced by work and disability, respectively. The coefficient of risk aversion is greater than one so that individuals are risk averse, which is an important assumption when studying the effect of SI on behaviour.[54]

Individual's Problem in the Labour Market

These features of individuals and the market environment define an individual's decision problem for each period in the labour market. Each period, individuals choose whether to participate in the labour market and earn employment income, $L_{it} \in \{0, 1\}$, or to apply for DI if eligible, $m_{it} \in \{0, 1\}$, to solve:

$$\max_{L,m} V_{it} = \mathbb{E}_t \Bigg(\sum_{s=t}^{T} \beta^{s-t} U(c_{is}, L_{is}; d_{is}) \Bigg| \Omega_t \Bigg), \tag{10}$$

$$\text{s.t. } c_{it} = \tau\big(W_{it} L_{it},\ DI_{it}\big) + SA(\tau\big(W_{it} L_{it},\ DI_{it}\big), d_{it}), \tag{11}$$

$$e_{it} = f(e_{it-1}, W_{it}, t). \tag{12}$$

Individuals decide to work or apply for DI in order to maximize their discounted lifetime utility, equation (10), subject to their budget constraint (11) and the evolution of their earnings index (12).[55] Utility from future periods is discounted by β. The expectation operator, $\mathbb{E}_t$, is conditional on the set Ω_t, which includes

[52] Variants of this specification for preferences are common in related studies, such as, Low and Pistaferri (2015), Michaud and Wiczer (2018), and Kostøl et al. (2019).

[53] Note the utility cost of work nets out any disutility from being on SA or DI.

[54] Risk aversion means individuals dislike uncertainty, which raises the relative value of insurance programs.

[55] The model assumes income is consumed each period, as there is no market for savings. Early-onset individuals have less scope for savings because of lower earnings. Moreover, this group has less incentive to save, given the higher generosity of the SI environment. As the research focus is education choices, the priority is to fit the earnings return to education. A related question is decomposing the value post-secondary into consumption or self-insurance, but this is outside the scope of this paper.

individual heterogeneity, $\{d_{i0}, a_i\}$, and time-varying state variables coming into the period, S_t. The state variables in a given period include current disability status, d_{it}, the current idiosyncratic shock to productivity, ϵ_{it}, the value of their earnings index from the previous period, e_{it-1}, and their eligibility for DI, $elidg_{it}$. The agent's expectation is over the sources of risk in my setting, which include disability risk, idiosyncratic productivity risk, the job arrival rate, and the job destruction rate.

The budget constraint is an equality under the assumed parametric structure on preferences, implying income from all sources is consumed. W_{it}, DI, and SA are the monetary values of labour earnings, DI benefits, and SA benefits, respectively. Labour earnings and DI benefits are subject to income taxes through the function, $\tau()$, representing the Canadian combined provincial and federal tax system.[56] An individual receives labour income when employed, $L_{it} = 1$, and they receive DI benefits if they are eligible, $elidg_{it} = 1$, have chosen to apply, $m_{it} = 1$, and are accepted to the program. The monetary value of SA benefits depends on the individual's income from other sources being below the poverty threshold, as described above.

Psychic Cost of Education

Education is a costly investment in terms of financial resources and utility and may be more costly in the presence of a disability. In the model, this is captured by the idiosyncratic psychic cost of post-secondary education, ψ_i. I normalize the psychic cost associated with low education to zero. The utility cost to education depends on initial disability d_0, plus an error capturing an idiosyncratic preference shock for education, $\epsilon_\psi \sim N(0, \sigma_{\epsilon^\psi})$.[57] The psychic cost of post-secondary education is

$$\psi_i = g_0 + g_1 d_{i0} + \epsilon_i^\psi. \tag{13}$$

5.3 Value Functions and Model Solution

I solve the model numerically via backward induction, as there is no analytical solution. The solution algorithm is straightforward, as each period's decisions and policy functions are conditional discrete choices. In the following, I suppress the individual's subscript, i, to simplify notation. Beginning with the terminal condition of retirement at age 65, I iterate backward, numerically approximating the value functions, characterizing the work decision and DI application decision at each age after eighteen as a function of

[56] The Canadian income tax system is a discrete set of tax rates and respective tax brackets. The tax parameters are calculated based on the weighted average of combined federal-provincial rates and brackets over the calendar years covered by my sample. For more details on the parameters of the tax and transfer system, refer to the Appendix.

[57] This cost helps rationalize education choices that are not consistent with ability sorting. This shock may include factors influencing the education decision that may be related to budgetary and funding differences, the intrinsic value to completing a post-secondary degree, as well as differences in preferences.

$S_t = \{d_t, \epsilon_t, e_{t-1}, elidg_t\}$.[58] Given the solution to the individual's labour market decisions, I solve the policy function for the education choice at age eighteen as a function of initial heterogeneity, $\{a, d_0, \psi\}$.

Retirement

Solving the model starts with the terminal condition, retirement. I assume that state variables remain fixed as soon as an individual retires, $S_t = S_{t+1} = \bar{S} = \{\bar{d}, \bar{\epsilon}, \bar{e}, el\bar{i}dg\}$. Individuals make no decisions in retirement. They receive utility from consuming their retirement income, which is known with certainty given their earnings index at the end of their working life.[59] I assume individuals expect retirement to last until age 75, after which they die with certainty. The value of retirement is

$$\begin{aligned} V_t^R(\bar{S}) &= u^N(c_t; \bar{d}\ ) + \beta V_{t+1}^R(\bar{S}) \\ \text{s.t. } c_t &= 5500 + 0.25\bar{e}. \end{aligned} \tag{14}$$

Before retirement, individuals can find themselves in one of three states in the labour market; working, not working and receiving SA, or not working and receiving DI. I consider the value functions and timing of choices for each state in turn, for ages less than 60 when individuals are not subject to retirement risk.

Value of Working

Given S_t, employed individuals earn flow utility from consuming after-tax work income and from SA at the beginning of the period. Shocks to productivity and disability then update to ϵ_{t+1} and d_{t+1} and the earnings index updates given their labour earnings. Individuals then face the job destruction rate, $\delta^{d_0,s}$, which places them out of work in the next period. If their job is not destroyed, individuals may choose to continue working or leave work. The value function for employed individuals is

$$\begin{aligned} V_t^E(S_t) =&\ u^W(c_t; d_t) + \beta E_t\Big[\delta^{d_0,s} V_{t+1}^U(S_{t+1}) + (1-\delta^{d_0,s})\max\big\{V_{t+1}^U(S_{t+1}), V_{t+1}^E(S_{t+1})\big\}\Big] \\ \text{s.t. } c_t &= \tau\big(W_t L_t,\ 0\big) + SA_t(\tau\big(W_t L_t,\ 0\big), d_t), \\ e_t &= f(e_{t-1}, W_t, t). \end{aligned} \tag{15}$$

[58] This approach to solving the life-cycle model is standard in finite horizon discrete choice dynamic programming models (Low et al., 2010; Low and Pistaferri, 2015). Additional details on the numerical solution can be found in the Appendix.

[59] The individual's contribution period ends at T^L so their earnings index remains constant after this time.

Value of Not Working and Receiving SA

While out of work, an individual receives flow utility from consuming SA income. Then, if eligible, they choose to apply for DI, $m_t = 1$, to become a beneficiary at the beginning of the next period. If applying, they are accepted with probability π^s. If accepted, their disability and productivity shocks update and their earnings index becomes fixed. If rejected, they do not receive a job offer and remain out of work for the next period. If the agent does not apply, $m_t = 0$, then their productivity and disability status update, and they receive a job offer with probability $\lambda^{d_0,s}$. If offered, they choose to accept and enter work the next period or to reject and remain out of work the next period. If the individual does not receive a job offer, they remain out of work for the next period. The value function for an unemployed individual at age t is

$$\begin{aligned} V_t^U(S_t) = u^N(c_t; d_t) + \beta\, \mathrm{E}_t \max_{m_t \in \{0,1\}} \Big[& m_t\big(\pi^s V_{t+1}^{DI}(S_{t+1}) + (1-\pi^s)V_{t+1}^U(S_{t+1}) - C_{app}^{d_0,s}\big) \\ & + (1-m_t)\Big(\lambda^{d_0,s} \max\big\{V_{t+1}^U(S_{t+1}), V_{t+1}^E(S_{t+1})\big\} + (1-\lambda^{d_0,s})V_{t+1}^U(S_{t+1})\Big)\Big] \\ \text{s.t. } c_t &= SA(0, d_t), \\ e_t &= f(e_{t-1}, 0, t). \end{aligned} \tag{16}$$

DI Beneficiary

I assume that individuals cannot work when receiving DI but can receive SA benefits simultaneously. Periods that the individual receives DI are not included in their contribution period. Therefore, their earnings index does not change when on DI. DI beneficiaries face the risk of reassessment of benefits, ρ. If benefits are not reassessed, the individual may or may not receive a job offer. If they receive an offer, work is added to their choice set. The value function for a DI recipient is

$$\begin{aligned} V_t^{DI}(S_t) = \; & u^N(c_t; d_t) + \beta E_t\Big[\rho V_{t+1}^U(S_{t+1}) + (1-\rho)\Big((1-\lambda^{d_0,s})\max\{V^U(S_{t+1}), V^{DI}(S_{t+1})\} \\ & + \lambda^{d_0,s}\max\{V^E(S_{t+1}), V^U(S_{t+1}), V^{DI}(S_{t+1})\}\Big)\Big] \\ \text{s.t. } c_t &= \tau\big(0,\ DI_t\big) + SA_t(\tau\big(0,\ DI_t,\big), d_t) \\ e_t &= e_{t-1}. \end{aligned} \tag{17}$$

Education Choice

The schooling decision is made at t=0 based on the expected value of each schooling level, $V_0(d_0, a, s)$. The value depends on initial disability status, the ability endowment, and education level. Individual i will choose to go to post-secondary if

$$V_0(d_0, a_i, s = 1) - V_0(d_{i0}, a_i, s = 0) - \psi_i \geq 0. \tag{18}$$

The inequality in equation (18) captures how early disability may influence educational investments by affecting these value functions. With a continuum of rational, forward-looking agents, there is a group on the margin of choosing higher education. For SI policy, the expected future recipiency of SA or DI is contained within the value functions for each education level. Therefore, any changes in the expected recipiency necessarily shift the group of individuals on the margin.

6 Model Estimation and Identification

In the first stage, a set of parameters is calibrated to values from the related literature or are estimated external to the life-cycle model. In the second stage, I estimate the remaining parameters with indirect inference, given the parameter values obtained in the first stage. The indirect inference method specifies an auxiliary model to capture key identifying moments in the data and then chooses the remaining model parameters to match these moments as closely as possible using data simulated from the model.

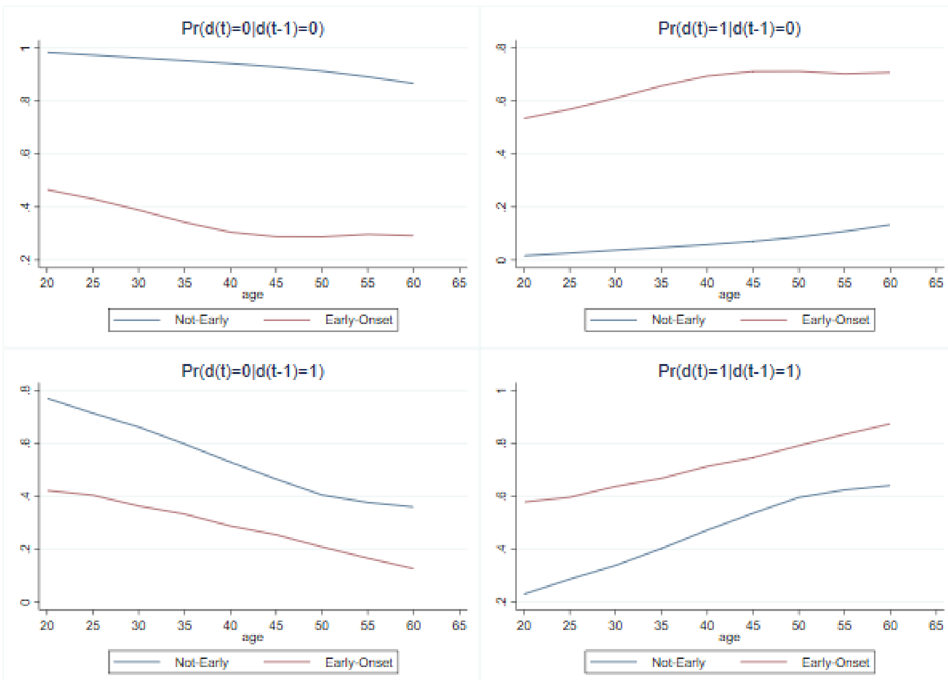
First, I set the coefficient of risk aversion to the value used in Low and Pistaferri (2015) and in Blundell et al. (2016), $\kappa = 1.5$. This ensures individuals are risk averse in the model. The discount factor is calibrated to $\beta = 0.9756$, the value used in Low and Pistaferri (2015).[60] Moreover, I take the value of the utility loss from disability in terms of consumption from Low and Pistaferri (2015), who set $\theta = -0.488$. The annual rate of reassessment and termination of DI benefits, ρ, is calibrated to 8%, which equals the rate of DI termination due to recovery as reported in The Canada Pension Plan Experience Study of Disability Beneficiaries in 2011 (Cheng et al., 2011).

[60] κ is in a comparable range as estimated in Attanasio et al. (1999), Attanasio and Weber (1995), and Banks and Brugiavini (2001). In Low and Pistaferri (2015), β reflects the annual discount factor from Gourinchas and Parker (2002) and Cagetti (2003).

6.0.1 Disability Risk

Under the assumption that disability risk is exogenous to the choices of agents in the model, I estimate a set of disability transition probabilities using observed transitions in the survey waves of LISA.[61] I obtain estimates of the transition probabilities, $\gamma_{k,l}^{t,d_0}$, by regressing an indicator of the joint event $\{d_{it} = k, d_{i,t-1} = l\}$ on a set of age dummies conditional on $d_{i,t-1} = l$ and d_0.[62] I use the predicted coefficients from this regression and smooth them by locally weighted regression to recover transition probabilities over the life cycle. The resulting disability transition probabilities over the life cycle are reported in Figure 1.

Figure 1: Disability Transition Probability Over the Life-Cycle



[a]

[a]Life-cycle transition probabilities are derived from 2012-2018 LISA survey waves. Transition probabilities are calculated for 5-year age bins, then smoothed with a LOWESS approach.

The top right panel of Figure 1 shows that the probability of inuring a disability shock increases with age and is much higher for early-onset individuals. This is consistent with a high rate of disability re-occurrence in adulthood for those with an early-onset disability. The bottom left panel in Figure 1 shows the likelihood of disability recovery decreases with age and is much lower for early-onset individuals. This is consistent with early-onset disabling conditions having a greater degree of persistence.

[61]Disability status is unobserved in the tax records. I implicitly assume the transition probabilities in the calendar years covered by LISA, 2012-2018, represent the calendar years covered by the tax records, 1989-2018.

[62]Age dummies reflect eight 5-year age windows, starting at age 25, and one age window from age 18-25.

6.0.2 Job arrival and destruction rates

The job arrival and job destruction rates in the model are set equal to the values reported in Table 2 of Section 4. The exogenous job offer arrival rate is estimated from the survey waves of LISA, which identifies individuals who are no longer with the same employer from the previous interview. This sub-sample is then asked why they are no longer with the same employer, and those reporting they were fired, laid off, or on a temporary contract that finished are flagged as exogenously losing their job. The job destruction rate is then adjusted to an annual rate.

The job arrival rate is estimated from a survey question that reports a respondent's labour market status for the previous 36 months. In this history of labour market states, an individual who is searching for work is distinct from one who is not in the labour market. The arrival rate is derived from the fraction of individuals who are searching in one month and employed in the next. Employment includes both part and full-time work. This recovers a monthly job arrival rate, which is then converted into an annual rate. These estimates of search frictions are held constant, assuming that the survey years are representative of frictions present in other calendar years that are covered in the tax data.

6.1 Indirect Inference

Estimation of the remaining structural parameters is achieved with indirect inference. Indirect inference is a simulation-based estimation technique used when an economic model's likelihood function is analytically intractable or too difficult to evaluate.[63] The main ingredient of indirect inference is an auxiliary model that captures key moments in the data that provide identifying information on the remaining structural parameters. Indirect inference chooses the economic model's parameters to make estimates from the auxiliary model using the observed data as close as possible to estimates from the auxiliary model using data simulated from the economic model. The observed data is an unbalanced panel, and I replicate censoring in the observed data by age, early disability status, and education level when calculating the moments from the simulated data.

The set of estimated parameters, $\hat{\Theta}$, are defined by

$$\hat{\Theta} = \underset{\Theta}{argmin}\left\{ \sum_{k=1}^{K} \left[(M^d_{kN} - M^m_{ks}(\theta))^2 / Var(M^d_{kN}) \right] \right\}, \tag{19}$$

[63]My framework is in the class of discrete choice models over a set of random utilities. These utilities are latent, but I observe the choices made by agents.

where the sum is over K moments, M_{kN}^d is the kth moment estimated over N observations, and $M_{ks}^m(\theta)$ is the kth simulated moment evaluated at parameter value Θ over s simulations. I weigh each mean squared difference using the variance of M_{kN}^d, $Var(M_{kN}^d)$.[64] An exception is moments capturing the distribution of post-secondary, for which I use an order of magnitude smaller than the variance of the data moment.[65] I have 201 moment conditions in the auxiliary model to estimate 32 structural parameters in the life-cycle model. I simulate the life-cycle decisions of 10,000 individuals (2 replications of 5,000 individuals).

6.1.1 Parameters of Earnings, Ability, and the Return to School.

I include a number of moments to identify the parameters of the earnings process. First, I estimate the coefficients from the following first differences linear model

$$\ln W_{it} - \ln W_{it-1} = \hat{\mu}_1^{d_0,s}(PE_{it} - PE_{it-1}) + \hat{\mu}_2^{d_0,s}\frac{(PE_{it}^2 - PE_{it-1}^2)}{100} + \hat{\phi}(d_{it}^* - d_{it-1}^*) + \hat{\xi}_{it}^{d_0,s}. \tag{20}$$

This model is estimated separately by d_0 and s. Notice that this model is similar to the sample analog of the first difference of the earnings process and provides identifying information to the coefficients on PE, PE^2, and ϕ. An important difference in this model is that d_{it}^* is an absorbing disability state upon onset. These parameters are bias estimates as the model does not correct for selection into employment. However, this model simulates this selection so as to match the selection bias present in the real data.[66]

With, $\hat{\mu}_1^{d_0,s}, \hat{\mu}_2^{d_0,s}$, and $\hat{\phi}$, I to compute

$$\hat{v^{d_0}}(a_i, s_i) = T_i^{-1}\sum_t \left(lnW_{it} - \hat{\mu}_1^{d_0,s}PE_{it} - \hat{\mu}_2^{d_0,s}\frac{PE_{it}^2}{100} - \hat{\phi}d_{it}^*\right), \tag{21}$$

where T_i is the number of years the individual has been observed working. Given the model's parametric assumptions, $v^{d_0}(a_i, s_i)$ is an estimate of individual human capital at labour market entrance, $h_s^{d_0}a_i + \xi_0$. I use the mean, variance, and the earnings cutoff for the first three quartiles of $\hat{v^{d_0}}(a_i, s_i)$ conditional on s and d_0 to identify $\sigma_{\xi_0}^2$, h_1^0, h_1^1, $\bar{a}_0$, $\bar{a}_1, \sigma_{a^0}^2$, and $\sigma_{a^0}^2$ relative to the normalization that $h_0^0 = h_0^1 = 1$.[67] To supplement these moments, I also compute the mean and variance of annual earnings in the first three periods in the labour market, conditional on s and d_0. The first three periods are affected by productivity

[64] I follow Blundell et al. (2016) is using $Var(M_{kN}^d)$ to as the weighting matrix, as the asymptotically optimal weight matrix has potentially poor small sample properties Altonji and Segal (1996).

[65] This decision helps to match the schooling distribution in the data, which have lower weight relative to the precise moments for life cycle earnings and employment. However, this decision comes at the expense of the efficiency of the parameter estimates.

[66] In many related studies, the earnings process is estimated exogenous to the rest of the model. However, in my application, the data limitations on disability made it difficult to do so.

[67] Note that difference in h_0^0 and h_0^1 are not separately identified from $\bar{a}_{d_0}$ and $\bar{a}_{d_1}$, respectively.

shocks the least, so these additional moments provide similar estimates of the same object.

The residuals from equation (20) are bias sample analogues of the individual productivity shocks, $\xi_{it}^{d_0,s}$. The difference is that earnings variation due to disability transitions is be contained in $\hat{\xi}_{it}^{d_0,s}$, since d^* is absorbing. I estimate the residual,

$$\hat{\xi}_{it}^{d_0,s} = (\ln W_{it} - \ln W_{it-1}) - \hat{\mu}_1^{d_0,s}(PE_{it} - PE_{it-1}) + \hat{\mu}_2^{d_0,s}\frac{(PE_{it}^2 - PE_{it-1}^2)}{100}) + \hat{\phi}(d_{it}^* - d_{it-1}^*), \qquad (22)$$

separately by s and d_0 and calculate it's group specific variance to pin down $\sigma^2_{\xi^{d_0,s}}$. Again, the moment calculated with the simulated data will replicate the bias in $\xi_{it}^{d_0,s}$.

To further help recover the true earnings profile over the life-cycle, match the first four quantiles of the lifetime earnings distribution and I use the OLS estimates from the following conditional regression: regression:

$$LnW_{it} = \beta_0^{d_0,s}t + \beta_1^{d_0,s}\frac{t^2}{100} + \beta_3^{d_0,s}d_{it}^* + u_{it}^{d_0,s}. \qquad (23)$$

Additionally, I match the variance of residuals from these models, $u_{it}^{d_0,s}$, which contain identifying information for productivity shocks in the model.

In estimating the parameters of the earnings process, I allow for measurement error. To separately identify productivity shocks from measurement error, I estimate the following model,

$$LnW_{it} = \beta_0 + \beta_1 t + \beta_2\frac{t^2}{100} + \beta_4 LnW_{it-1} + u_{it} \qquad (24)$$

The variance of measurement error is identified off the variance and autocovariance of the residual from this model. This implies the assumption that measurement error is independent of disability status and schooling level.

Parameters of Disability Insurance

DI applications are unobserved in my data. I address this issue with a chosen set of moments that pertain to flows onto DI and rates of DI receipt. Given a set of the model's structural parameters, I can simulate DI applications and the resulting moments, which are reflective of the decision to apply for DI. Comparing the fit of the moments that relate to DI using the real data and simulated data is a useful check of how well the model predicts the decision to apply for DI. I partition the sample by s, d_0, and match the conditional rate of DI receipt, the conditional composition of DI recipients, and the conditional flow rates into DI. These moments are similar to those used in Low and Pistaferri (2015) and relate directly to the probability of successful application given the eligibility parameters of the program. That is, If the parameters governing DI are such that there is a higher probability of acceptance for a given disability severity and schooling level, then this would lead to a higher flow into DI and a larger proportion of recipients to DI for said disability severity and education level.

I also match OLS estimates from the following two conditional models:

$$
\begin{aligned}
\mathbb{1}(DI_{it}) &= \beta_0^{s,d_0} + \beta_1^{s,d_0} t + \beta_2^{s,d_0} t^2 + \beta_3^{s,d_0} t^3, \text{ and} \\
\mathbb{1}(DI_{it} = 1 \ \& \ DI_{it-1} = 0) &= \beta_0^{s,d_0} + \beta_1^{s,d_0} t + \beta_2^{s,d_0} t^2 + \beta_3^{s,d_0} t^3,
\end{aligned} \tag{25}
$$

where $\mathbb{1}(DI_{it})$ is an indicator variable that equals one if the individuals in on DI in period t and $\mathbb{1}(DI_{it} = 1 \ \& \ DI_{it-1} = 0)$ is an indicator variable that equals one if the individuals flowed onto DI in period t. These moments help the model fit life-cycle trends in DI application and enrolment, which are mostly zero in early life, then grow at an increasing rate after age 45 for all groups.

Last, to identify the utility cost of DI application, I match the unemployment rate and the cutoffs of the first five quantiles of observed earnings two periods prior to benefit receipt. The identifying argument is that the distribution of pre-application earnings will vary with the utility cost of the application. The utility cost of the DI application will be adjusted to match the pre-application earnings distribution observed in the data.

Utility Cost of Working.

To identify the utility cost of working, I match employment rates over the life-cycle conditional on (s, d^*, d_0) and age greater than or equal to 45 or less than 45. In addition, I match the flows into and out of employment by the same conditioning variables. These moments are informative of the utility cost of working, η. The idea is that, given δ^{d_0}, young workers have a lower return to working as they accumulate potential experience, which influences their employment decisions for a given value of η. Hence, the utility cost of working is identified by matching variation in employment rates and flows by age groups, as variation in the financial returns of working is identified off of the parameters of the earnings process and policy environment.

The probability of receiving SA-D benefits is identified from the employment rates of individuals who currently have and do not have a disability. Employment rates of individuals who currently have a disability reflect the probability they were approved for the higher SA-D threshold, altering their relative value of working. Hence, this probability is identified through matching employment rates conditional on current disability status, given the identification of the parameters of DI and the earnings process.

Parameters of Psychic Cost to School

The parameters governing the psychic cost to school are identified off of education distributions and a linear probability model of schooling on d_0, $\hat{v^{d_0}}(a_i, s_i)$, and an intercept. The regression coefficient and distribution help pin down the mean schooling by d_0. The schooling decision depends on endowed ability, which depends on d_0. In the probit model, I control for $\hat{v^{d_0}}(a_i, s_i)$, which is a direct estimate of ability plus a shock that is independent of psychic costs. Hence, the variance of the residual in this regression identifies the variance of psychic costs.

7 Empirical Results

The remaining sections review the estimation results and implications of the structural model. I first interpret the estimated parameters and detail the fit of the estimated model to moment counterparts in the data. I then use the estimated model to investigate features that contribute to the gap in education between early-onset individuals and their non-disabled counterparts. Lastly, I conduct counterfactual experiments that reform the policy environment and analyze the effects on education investments, life-cycle behaviour, and welfare.

First, I discuss and interpret the structural parameters estimated via indirect inference. Table 5 reports the estimated parameters of individual heterogeneity and the annual earnings process. At labour market entry, an individual's earnings depend on their endowed ability, education level, and a stock of human capital that is unrelated to schooling, ϵ_0. The mean of ϵ_0, presented in the first row of the bottom panel of Table 5, implies that 26% of initial human capital is unrelated to education for early-onset individuals and 20% for not early disabled individuals. The part of initial human capital augmented by schooling, endowed ability shown in the first row of the upper panel, is the predominant component of initial earnings. Mean ability is lower for early-onset individuals, implying a -$3800 difference in initial earnings at labour market entry relative to individuals that are not early disabled. Ability endowments are more volatile for early-onset individuals. This is consistent with the disruption of skill accumulation before age eighteen, resulting in a

Table 5: Estimates of Parameters for Individual Heterogeneity and the Earnings Process.

		Not Early Disabled		Early-Onset	
Conditional Parameters		Low Education	Post-Secondary	Low Education	Post-Secondary
Mean Ability	$\bar{a}^{d_0}$	7.3476 (0.0003)		6.8510 (0.1318)	
Variance of Ability	$\sigma^2_{a^{d_0}}$	0.1505 (0.4087)		0.4032 (0.0622)	
Return to Post-Secondary	h^{d_0}	1.0420 (0.0000)		1.0307 (0.0925)	
Return to Potential Experience	μ_1^{s,d_0}	0.1052 (0.0000)	0.1231 (0.0001)	0.0949 (0.0003)	0.1225 (0.0000)
$\frac{\textit{Return to Potential Experience}^2}{100}$	μ_2^{s,d_0}	-0.2297 (0.0000)	-0.2557 (0.0001)	-0.1727 (0.0003)	-0.2629 (0.0015)
Productivity Shock	μ_1^{s,d_0}	0.0128 (0.0044)	0.0065 (0.0000)	0.0212 (0.0000)	0.0158 (0.0000)
Unconditional Parameters					
Earning Penalty of Disability	ϕ	-0.0345 (0.0000)			
Mean Initial	$\bar{\epsilon}_0$	1.8377 (0.0018)			
Variance	$\sigma^2_{\epsilon_0}$	0.1091 (0.0007)			

Standard errors are in parenthesis below. These are calculated using the formula for the asymptotic variance, corrected for simulation error, from Gourieroux, Monfort, and Renault (1993).

greater range of human capital at the end of high school for early-onset individuals.[68]

Row three shows that attending post-secondary scales initial ability by 3.1% for early-onset individuals and 4.2% for not early disabled individuals. This result is consistent with disabilities disrupting the efficiency of human capital accumulation during post-secondary schooling. Row 4 shows that the efficiency of labour market human capital accumulation through potential experience, μ_1^{s,d_0}, is greater with post-secondary education and lower for early-onset individuals, conditional on education. Earnings are much more volatile for early-onset individuals, and earnings volatility decreases with education. In row 6 we see that the onset of a disability in the labour market results in a 3.45% penalty to annual earnings.

The remaining structural parameters are reported in Table 6. First, the utility cost of working, η, equates to approximately 8% of annual consumption. The likelihood of acceptance to DI, shown in the third

[68]For instance, an early-onset disability may create barriers that drastically disrupt skill accumulation for some, and others may be able to easily accommodate their disability.

Table 6: Estimates of Remaining Parameters

Description	Parameter	Estimate
Utility Cost of Work	η	-0.0784
		(0.0073)
Variance of Measurement error	$\sigma^2_{\epsilon_{ME}}$	0.0797
		(0.0000)
Policy Parameters		
DI Acceptance Probability for s=0	π_0	0.4645
		(0.3669)
DI Acceptance Probability for s=1	π_1	0.4088
		(0.2278)
Application Cost of DI (Utility) for s=0	di app cost s0	0.0001
		(0.0000)
Application Cost of DI (Utility)for s=1	di app cost s1	0.0001
		(0.0000)
Application Cost of DI (Utility)for $d_0 = 1$	di app cost d0 = 0	0.0001
		(0.0010)
SA Disability Benefits Acceptance Probability	π^{SA}	0.8149
		(0.0024)
Psychic Cost Parameters		
Average Psychic Cost of Post-Secondary	g_0	0.0058
		(0.0008)
Average Psychic Cost of Post-Secondary for $d_0 = 1$	g_1	0.0027
		(0.0008)
Variance of Psychic Cost of Post-Secondary	$\sigma^2_{\epsilon_\psi}$	0.0208
		(0.0114)

Standard errors are in parenthesis below. These are calculated using the formula for the asymptotic variance, corrected for simulation error, from Gourieroux, Monfort, and Renault (1993).

and fourth rows, is modestly lower for applicants with post-secondary.[69] The unconditional acceptance rate of 43% during the 2014-2015 fiscal year is remarkably similar to the simulated unconditional acceptance rate of 43.9% implied by the model (Office of the Auditor General of Canada, 2015). The acceptance rate is not targeted in estimation and serves as external validation of the model's DI program approximating CPP-D. Applications to DI impose a cost to utility, which differs by early disability status. This cost depends on the value of consumption in the outside option to DI application. To illustrate, for an outside option of consuming $10,000, the utility cost of application is equivalent to approximately $85 for all applicants, plus an additional $39 for individuals not initially disabled $d_0 = 0$. Lastly, 81% of individuals with a disability

[69] I found no substantial difference in the likelihood of DI acceptance conditional on s when allowing this parameter to vary by applicants aged < 45 vs ≥ 45 in an alternate model specification.

are accepted for SA-D benefits.

The parameters for the psychic cost to schooling are presented bottom three rows. Individuals in the model incur a positive psychic cost when going to post-secondary, which is larger with an early-onset disability. The utility cost for all individuals equates to an average reduction in yearly consumption of $634 each year in school. Individuals with an early-onset disability have an additional utility cost equal to an $85 per year reduction in average consumption during school.

7.1 Model Fit

Table 7: Simulated Likelihood of Post-Secondary Relative to Likelihood in the Actual Data.

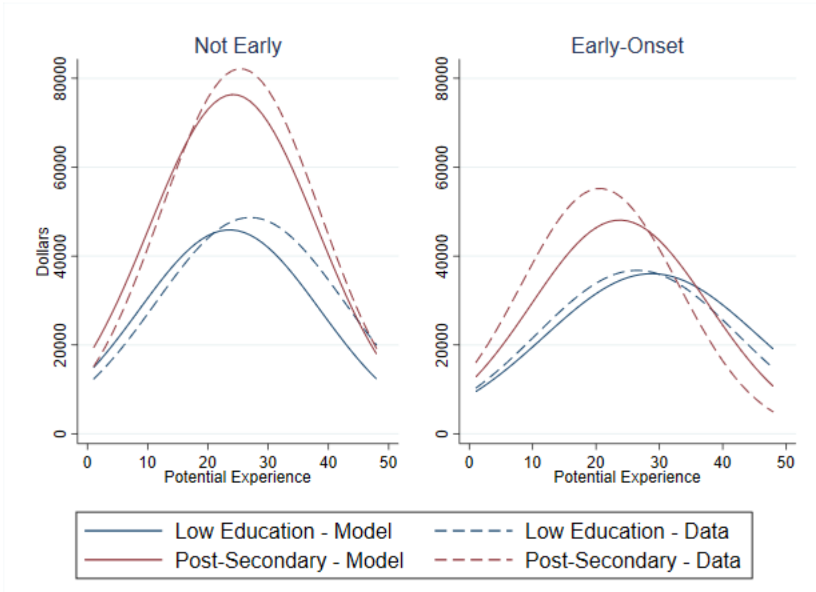
	Moment	Sim
Early-Onset	0.4600	0.4667
	(0.0370)	
Not Early Disabled	0.6400	0.6371
	(0.0120)	

Notes: Standard errors of moment in data are in parenthesis below the statistic.

Next, I compare the fit of the estimated model relative to its moment counterparts in the data. As it can be difficult to interpret the values of estimated parameters in a large structural model, contrasting the true moments with moments calculated using data simulated from the model helps to validate the parameters with their interpretation. First, Table 7 shows the rate of post-secondary education attainment by early disability status. The model replicates the education choice very well for both the early-onset and not early disabled groups. Additionally, the model's simulated gap in educational attainment, 17 percentage points, is very similar to the 18 percentage point gap observed in the data.

A prominent component of the return to education is the financial gain that results from investing in human capital. Figure 2 shows the predicted life-cycle annual earnings profile separately by early disability status and education level. The red lines are conditional on post-secondary, and the blue linear are conditional on low education. The left figure pertains to the not early disabled group, where the solid line is from the actual data, and the dotted line is from data simulated with the model. The model recovers a similar earnings profile with respect to potential experience as in the data. Moreover, the data recovers a similar earnings premium to post-secondary schooling. The right figure pertains to the early-onset group. Again, the model recovers a similar life-cycle earnings profile for each education group. For this group, the education premium

Figure 2: Model Fit: Life-cycle Earnings



Note: Predicted profile of life-cycle annual earnings are derived from estimating equation 20. Predicted earnings from true data are shown by dotted lines and predicted earnings using simulated data are shown by solid lines.

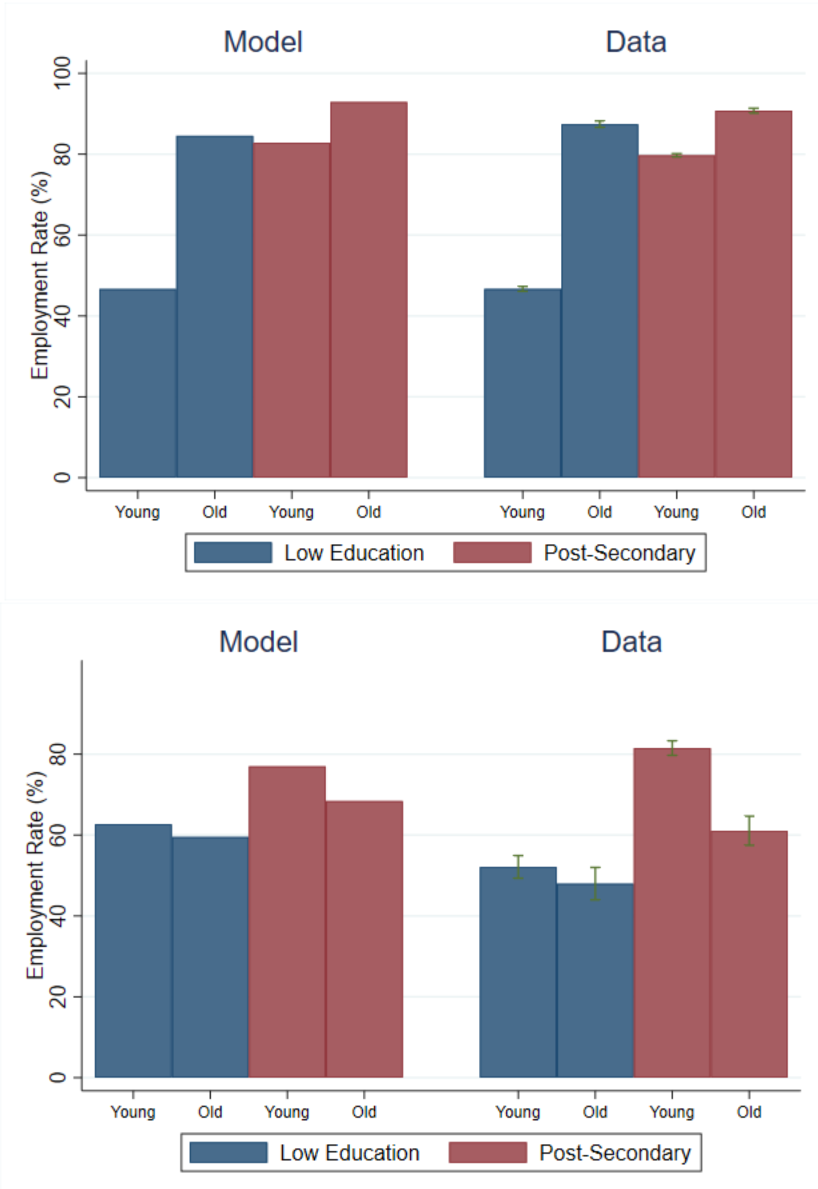
is attenuated early in life. However, the difference by early disability status within each education level is very similar that observed in the actual data.

Annual earnings represent the return to working relative to the outside option of not working and collecting government transfers. Figure 3 reports aggregate employment rates separately by early disability status, education level, and age groupings. Again, the red bars are conditional on post-secondary education, and the blue bars are conditional on low education. The left graph shows rates simulated from the model, and the right graph shows rates from the actual data. First, the model reproduces the increase in employment with age for the not early disabled group, shown in the top figure, within each education level. Additionally, the model reproduces a higher employment rate by education very well. The lower figure reports the same for the early-onset group. Again, the model reproduces a reduction in aggregate employment by age groups and a positive relationship between education and employment.[70]

Lastly, Table 8 presents rates DI by early disability status and education level over all years in the labour market. The estimated model matches the rate of DI very closely. Early-onset individuals have the highest percentage of their population on DI. Moreover, I recover a positive correlation between the rate of DI and age. The rate of DI is nearly zero before age 45, then grows at an increasing rate as individuals

[70]Employment rate for the early-onset group is over-predicted by about 10% in the model, which is related to the attenuation in annual earnings for early-onset individuals with post-secondary.

Figure 3: Model Fit: Aggregate Employment



Note: Graph is derived from employment statistics in Table 14 in Appendix. Employment rates for not early disabled individuals is on top and early-onset individuals is on bottom.

approach retirement.

Table 8: Simulated Rate of Employment and DI over the Life-Cycle Relative to Rates in the Actual Data

Rate of DI	Not Early Disabled Data	Sim	Early-Onset Data	Sim
Low Education	0.0238	0.0250	0.0396	0.0313
	(0.0010)		(0.0050)	
Post-Secondary	0.0085	0.0013	0.0407	0.0350
	(0.0004)		(0.0040)	

Notes: Standard errors of moment in data are in parenthesis below the statistic from the actual data.

8 Counterfactual Exercises

Next, I use the estimated model to decompose the observed education gap between early-onset individuals and their not early disabled counterparts. This decomposition motivates the role of SI policy on education investments relative to other model features that differ by early disability status. I then analyze the effect of counterfactual reforms to SI policy on education choices, life-cycle behaviour, and individual welfare.

8.1 Decomposing the Education Gap

The model predicts an education gap of 17 %-points. I use the estimated model to determine the most important features for this gap. The structural model specifies sets of parameters that differ by d_0, and I group them into a coarser set of contributing factors. The education choice hinges on one's expectation of the lifetime value of each level, which are functions of these parameters. To better understand the role of SI policy on education choices, it is useful to compare the role of other determinants of education investments between early-onset and not early disabled individuals.

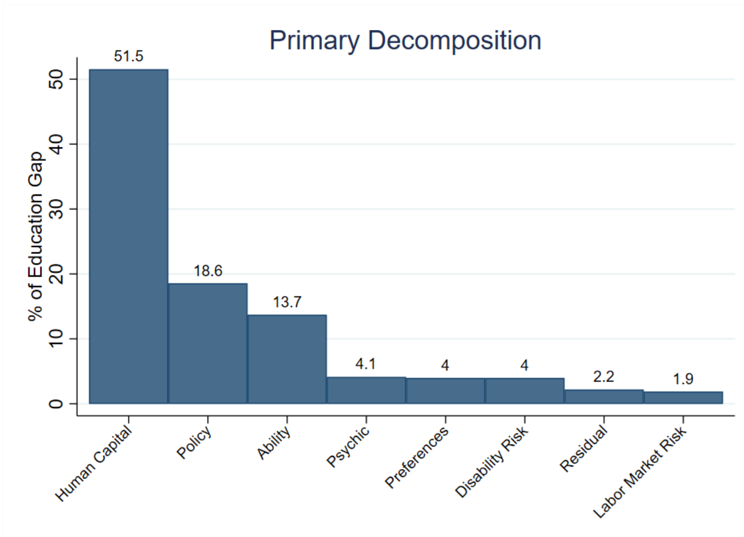
This section analyzes the drivers of the estimated education gap, hereby referred to as the baseline gap. To compare the relative contribution of different sets of parameters, I decompose the education gap using a version of a Shapley Decomposition (Shorrocks et al., 2013). This decomposition recovers the expected marginal contribution of a factor to the overall education gap. The contribution of a factor to the education gap is calculated through a process of equalizing subsets of parameters that differ by d_0 to their estimated value for not early disabled individuals. I then resolve the model and calculate the change in the education gap. I do this for all possible subsets of factors eliminated and average them based on their frequency in the set of all ordered elimination sequences.[71]

[71] For a step-by-step breakdown of this procedure, see (Shorrocks et al., 2013)

This decomposition method has three desirable features. First, the decomposition is exact, meaning each factor's contribution sums to equal the baseline gap.[72] Second, this decomposition is symmetric, which occurs when the estimation of each factor's contribution is path independent to the order estimating other factors' contributions.[73] Lastly, this method accommodates hierarchical structures within the contributing factors. I conduct a primary decomposition, where the contribution of the primary factors can be decomposed into the contribution of a set of secondary factors. I consider seven primary factors, three of which have a secondary structure.

Figure 4 plots the results from the primary decomposition. From left to right, the seven primary factors

Figure 4: Decomposing the Education Gap: Primary Decomposition



Note: Graph displays results of an Shapely decomposition on the 17% post-secondary education gap between early-onset and not early disabled individuals implied by the estimated model. Each bar represents the percentage contribution of the respective factor to the overall gap.

are human capital, policy, initial ability distribution, the psychic cost to schooling, the effect of disability on preferences, differences in disability risk for early-onset, and labour market risk. The residual category arises from simulation error. Human capital combines parametric differences in the return to post-secondary, h^{d_0}, the earnings return to potential experience, $\mu^{d_0,s}$, and the direct effect of disability, ϕ.

The human capital factor has the greatest contribution to the baseline education gap. Early-onset disability negatively impacts the accumulation of and return to human capital. Differences in the distribution of initial ability are the third-largest contributor to the gap, accounting for a 2.33 percentage point

[72]An example of a non-exact decomposition is to set the contribution of each factor according to turning that factor off with all others turned on. Results for this alternate type of decomposition can be found in Appendix.

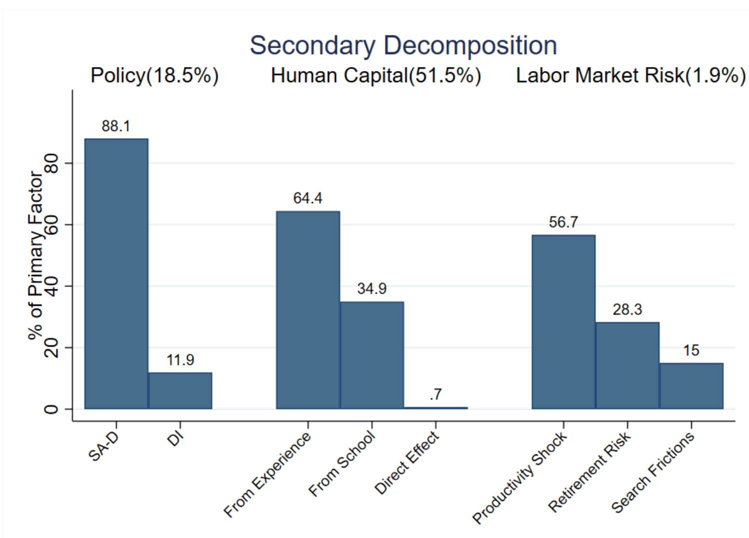
[73]An example of a path-dependent decomposition is to shut factors down in sequence.

difference in post-secondary completion. Early-onset disability interferes with the development of productive skills before age eighteen, resulting in lower starting average ability. Combined, human capital and ability account for two-thirds of the education gap. However, there is ample room for other factors, notably policy, to affect the education investments of early-onset individuals. This is missed in competitive settings where the interaction between health and ability drives education choice.

The second-largest contributor to the gap, labour market policy, results in 3.1 percentage points lower post-secondary completion for early-onset individuals. This factor combines both the DI and SA policies. With an early-onset disability, individuals have a better outside option to work, resulting in a greater relative value to low education. This result formalizes a mapping from SA and DI policies to pre-entry education choices.

Psychic costs, g_1, the utility cost of disability, θ, and the difference in disability risk all account for 0.7 percentage points less post-secondary completion for early-onset individuals. Interestingly, labour market risks, which combine job destruction, job arrival, and retirement risks, are the least predictive of education choice. These are explored further in the secondary decomposition.

Figure 5: Decomposing the Education Gap: Secondary Decomposition



Note: Graph displays results of an Owen decomposition on three primary factors from the Shapely decomposition of post-secondary education gap between early-onset and not early disabled individuals implied by the estimated model. Each bar represents the percentage contribution of the respective factor to the respective primary factor.

Figure 5 reports the results from the secondary decomposition. The primary factors for policy, human capital, and labour market risk have a secondary structure. First, the education-specific returns to experience account for two-thirds of the human capital factor. Early-onset disability impedes the accumulation of skills through work. The increase in ability from post-secondary accounts for one-third of the human capital factor.

Within the policy factor, SA-D is the main contributor to the education gap, accounting for 90%. This program, which raises the outside option of work, is especially valuable for early-onset individuals who face greater adversity in the labour market. DI has minor effects on education for two main reasons. First, people generally flow onto DI in the second half of working life when their earnings index is large enough and disability risk is higher. The option value of DI becomes heavily discounted in people's expectations when choosing education. Second, because the value of DI is one-to-one with the earnings index, individuals without an early-onset disability are more likely to have a high option value for this program when expecting to incur a disability shock. Hence, the non-early group is relatively more sensitive to this program than to SA.[74]

Labour market risks do not have a substantial effect on education. Within this factor, the distribution of idiosyncratic productivity shocks has the greatest effect on education. Search frictions and retirement risk have trivial effects on education. While the likelihood of finding and maintaining work is lower in absolute terms for early-onset, the relative difference in these risks is greater across education. Hence, early-onset has a greater return to education in their expectation of finding and holding a job relative to those not early disabled.

8.2 Counterfactual Policy Experiments

The decomposition exercise gives intuition into the factors driving the education gap. In this last section, I evaluate a set of counterfactual reforms to the policy environment on education choice, individual welfare, and government costs. To motivate, a policymaker, when considering the education gap, likely has the intuition that people are generally more productive when educated, and more productive people are less costly to support for the government. Hence, a reasonable policy objective is to increase education for the early-onset group to promote financial independence.

The decomposition exercise identifies various areas that policymakers can target to reduce this gap. However, a policy reform targeting one contributing factor may differ in cost and the marginal group affected compared to another. Hence, policies will vary in cost to the government and the welfare of individuals. An approach that facilitates comparison across policies is fixing a policy objective and comparing alternate

[74] This result is shown in the ceteris paribus decomposition in the Appendix.

policies that achieve said objective. I fix the policy objective of increasing post-secondary completion of early-onset individuals by one percentage points (from 46.7% to 47.7%). I then solve how this objective is achieved under different policy reforms, comparing the net costs and benefits to the government and the impact on individual welfare.

The first policy reform addresses the disincentives created by labour market policy by reducing SA-D by $1185. First, the previous section found SA-D to be an important contributor to the gap. Reducing this safety net will push people to self-insure their income through education. On the other side, many criticisms that welfare in Canada doesn't provide sufficient support to offset the cost of living. This counterfactual captures the insurance-incentive trade-off SI policies.

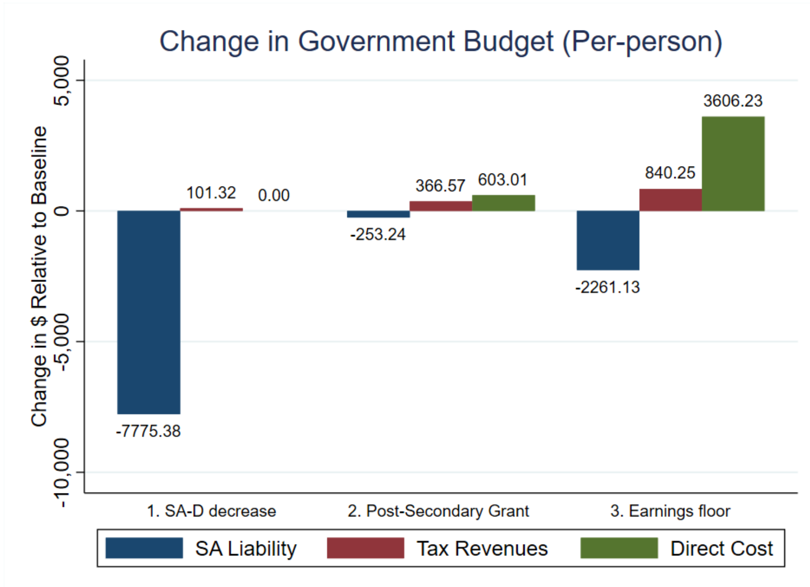
Second, I consider a reform that addresses the psychic cost of post-secondary. To offset the psychic cost, I calculate the value of a subsidy, which increases consumption during years of post-secondary for early-onset. I found a per year subsidy of $423 achieved a 1% increase. This policy targets individuals that might otherwise benefit from more education but choose not to do so because of a negative draw for their psychic cost.

The last counterfactual policy supplements income in the first three years of the labour market, when individual earnings are lowest. The reform provides an earnings floor of $12978 in the first three years after post-secondary for early-onset individuals. This policy targets individuals with low starting earnings (i.e., ability) and helps them gain a few years of experience to have better job opportunities when the subsidy expires.

Figure 6 summarizes the change in government revenue and liability relative to the baseline for each policy reform. The bars are the difference in average lifetime present value dollars per person. The blue bar shows the change in SA liability, the red bar indicates the change in tax revenues, and the green bar shows the direct cost of implementing each policy. First, reducing SA-D, unsurprisingly, results in substantial savings for governments. Although, removing this source of assistance has small increases in tax revenues. The schooling grant is slightly revenue positive. Individuals that switch to post-secondary in the counterfactual environment offset the cost of the subsidy with lower SA dependence and increased tax revenues for the government. The wage floor has large effects but is very costly to implement. This is driven by a large reduction in SA liability in the first three periods, which is replaced by the increased liability of providing the wage floor.

Finally, Figure 7 shows the difference in ex-ante individual welfare in each counterfactual policy environment relative to the baseline. I calculate the individual's average ex-ante willingness to pay (WTP) to

Figure 6: Effect of Counterfactual Policy Reform on the Change in Government Revenue and Liability.



Note: Figure displays the change in SA liability for government (blue), government tax revenues (red), and the direct cost of a policy reform (green) relative to the baseline scenario. Units are present value per-person lifetime dollars.

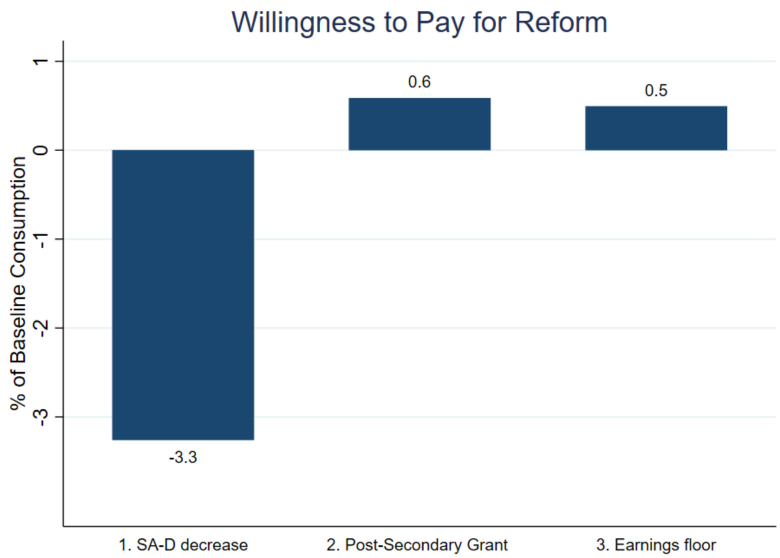
implement the policy change in terms of a proportional reduction to current and future consumption.[75] This measure is advantageous for welfare analysis as it is non-distortionary in the sense that it is equivalent to directly extracting utility from individuals. Additionally, this measure of WTP has an intuitive interpretation, representing the expected present value of the stream of consumption an individual is willing to forgo in the baseline scenario to live in the counterfactual policy environment.

Unsurprisingly, taking resources away from people at the left tail of the income distribution greatly reduces their welfare.[76] Early-onset individuals are willing to forgo 3.3% of their lifetime consumption stream in the baseline to retain their SA-D benefits. Hence, removing the work disincentives from SA-D has a considerable trade-off in terms of welfare for early-onset individuals, who greatly value this program. The welfare implications of the other two policies are very similar. The subsidy was revenue neutral for the government and is preferable to the wage floor along these two dimensions.

[75] The WTP is calculated as $WTP = (\frac{EV_{baseline}}{EV_{reform}})^{1-\kappa} - 1$.

[76] Nonlinearity of preferences means the marginal utility is greater at lower levels of consumption.

Figure 7: Individuals Willingness to Pay for Counterfactual Policy Reform



Note: Measure of willingness to pay is interpreted as how much (%) of the baseline stream of lifetime consumption willing to forgo to have this policy reform.

9 Conclusion

An early-onset disability can impose substantial disadvantages that persist throughout one's life. The effect of an early-onset disability can be mitigated through education investments. However, the incentives to invest in education depend on a number of factors, notably SI policy. In this paper, I build and estimate a structural model to analyze the relative importance of SI policy, the financial returns to education, and labour market risks on observed schooling investments of early-onset individuals. This paper gives insight into the many ways an early-onset disability influences education choices and analyzes the role of policy in improving the welfare and outcomes of this population.

The decomposition exercise finds the difference in the financial returns to schooling by early disability status plays the most considerable role in the education gap. The additional resources in SA programs for individuals with a disability also play an important role, accounting for 15.6% of the education gap between early-onset and not early disabled individuals. This result is because the expected value of these programs reduces the return to investing in education by raising the relative value of the outside option of working. DI has a trivial role for education investments, despite evidence that the expected value of this policy depends on one's schooling.

The policy reforms find that increasing the value of SA improves individual welfare but increases the moral hazard of the programs with respect to education investments and employment. Increasing the value of DI has trace effects on education. Reallocating resources from SA to DI reduces moral hazard of the policies, but at the expense of the individual's welfare. Instead, subsidizing consumption during schooling incentivizes education investment and increases employment, earnings, and consumption. This is due to people being more productive on average and having a higher return to work. Moreover, this policy helps pay for itself as more productive individuals create additional tax revenues and lower dependence on SI programs. Thus gains can be made at the individual level if governments focus policy efforts on enhancing individuals' economic independence.

References

Acemoglu, D. and Angrist, J. D. (2001). Consequences of employment protection? the case of the americans with disabilities act. *Journal of Political Economy*, 109(5):915–957.

Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature*, 56(4):1360–1446.

Altonji, J. G. and Segal, L. M. (1996). Small-sample bias in gmm estimation of covariance structures. *Journal of Business and Economic Statistics*, 14(3):353–366.

Ameri, M., Schur, L., Adya, M., Bentley, F. S., McKay, P., and Kruse, D. (2018). The disability employment puzzle: A field experiment on employer hiring behavior. *ILR Review*, 71(2):329–364.

Attanasio, O. P., Banks, J., Meghir, C., and Weber, G. (1999). Humps and bumps in lifetime consumption. *Journal of Business & Economic Statistics*, 17(1):22–35.

Attanasio, O. P. and Weber, G. (1995). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of political Economy*, 103(6):1121–1157.

Autor, D. H. and Duggan, M. G. (2006). The growth in the social security disability rolls: a fiscal crisis unfolding. *Journal of Economic Perspectives*, 20(3):71–96.

Baker, M., Stabile, M., and Deri, C. (2004). What do self-reported, objective, measures of health measure? *Journal of human Resources*, 39(4):1067–1093.

Banks, James, R. B. and Brugiavini, A. (2001). Risk pooling, precautionary saving and consumption growth. *The Review of Economic Studies*, 68(4):757–779.

Béland, D. and Daigneault, P.-M. (2015). The state of provincial social assistance in canada.

Benítez-Silva, H., Buchinsky, M., Man Chan, H., Cheidvasser, S., and Rust, J. (2004). How large is the bias in self-reported disability? *Journal of Applied Econometrics*, 19(6):649–670.

Black, N., Johnston, D. W., Shields, M. A., and Suziedelyte, A. (2017). Who provides inconsistent reports of their health status? the importance of age, cognitive ability and socioeconomic status. *Social Science and Medicine*, 191:9–18.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Bobba, M., Flabbi, L., Levy, S., and Tejada, M. (2021). Labor market search, informality, and on-the-job human capital accumulation. *Journal of Econometrics*, 223(2):433–453.

Bound, J. (1989). The health and earnings of rejected disability insurance applicants.

Bound, J. and Burkhauser, R. V. (1999). Economic analysis of transfer programs targeted on people with disabilities. *Handbook of Labor Economics*, 3:3417–3528.

Bound, J., Stinebrickner, T., and Waidmann, T. (2010). Health, economic resources and the work decisions of older men. *Journal of Econometrics*, 156(1):106–129.

Burkhauser, R. V., Daly, M. C., Houtenville, A. J., and Nargis, N. (2002). Self-reported work-limitation data: What they can and cannot tell us. *Demography*, 39(3):541–555.

Burkhauser, R. V., Haveman, R. H., and Wolfe, B. L. (1993). How people with disabilities fare when public policies change. *Journal of Policy Analysis and Management*, 12(2):251–269.

Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. *Journal of Business and Economic Statistics*, 21(3):339–353.

Carneiro, P., Hansen, K. T., and Heckman, J. J. (2003). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college.

Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of health economics*, 24(2):365–389.

Case, A. and Paxson, C. (2010). Causes and consequences of early-life health. *Demography*, 47(1):S65–S85.

Cheng, Y., Dontigny, P., Guimond, A., Harrel, S., Lemire, F., Montambeault, M., Pommainville, L.-M., and Van Le Tran, C. (2011). *Canada pension plan: Experience study of disability beneficiaries*. Office of the Chief Actuary.

Cossette, L. and Duclos, E. (2002). A profile of disability in canada, 2001.

Cunha, F., Heckman, J., and Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings. *oxford Economic papers*, 57(2):191–261.

Currie, J. (2009). Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development? *Journal of Economic Literature*, 47(1):87–122.

Currie, J. and Almond, D. (2011). Human capital development before age five. In *Handbook of labor economics*, volume 4, pages 1315–1486. Elsevier.

Cutler, D. M., Landrum, M. B., Stewart, K. A., et al. (2006). How do the better educated do it? socioeconomic status and the ability to cope with underlying impairment. *Developments in the Economics of Aging*, page 203.

David, H. (2011). The unsustainable rise of the disability rolls in the united states: Causes, consequences, and policy options.

Deshpande, M. and Dizon-Ross, R. (2023). The (lack of) anticipatory effects of the social safety net on human capital investment. Technical report, National Bureau of Economic Research.

Dixon, K. A., Kruse, D., and Van Horn, C. E. (2003). *Restricted access: A survey of employers about people with disabilities and lowering barriers to work*. John J. Heldrich Center for Workforce Development, Rutgers.

Druckman, J. N., Levy, J., and Sands, N. (2021). Bias in education disability accommodations. *Economics of Education Review*, 85:102176.

Employment and Social Development Canada (2016). Social assistance statistical report: 2009-2013. `https://www.canada.ca/en/employment-social-development/services/social-assistance/reports/statistical-2009-2013.html`.

Employment and Social Development Canada (2018). Annual report of the canada pension plan for fiscal year 2017 to 2018. `https://www.canada.ca/en/employment-social-development/programs/pensions/reports/annual-2018.html`.

Flinn, C. and Mullins, J. (2015). Labor market search and schooling investment. *International Economic Review*, 56(2):359–398.

Gallipoli, G. and Turner, L. (2009). Household responses to individual shocks: Disability and labor supply. Technical report.

Golosov, M. and Tsyvinski, A. (2006). Designing optimal disability insurance: A case for asset testing. *Journal of political Economy*, 114(2):257–279.

Gourinchas, P. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.

Government of Canada (2022). Canada pension plan adjudication framework. `https://www.canada.ca/en/services/benefits/publicpensions/cpp/cpp-disability-benefit/medical-professionals/framework.html#h2.2`.

Government of Canada (2023). Cpp contribution rates, maximums and exemptions.

Haveman, R. and Wolfe, B. (2000). The economics of disability and disability policy. *Handbook of Health Economics*, 1:995–1051.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the national Academy of Sciences*, 104(33):13250–13255.

Hillel, I. et al. (2020). Holes in the social safety net: Poverty, inequality and social assistance in canada.

Kellogg, M. D. (2021). *Spousal Labor Supply and the Welfare Implications of Disability Insurance Reform*. PhD thesis, The University of Chicago.

Kitao, S. (2014). A life-cycle model of unemployment and disability insurance. *Journal of Monetary Economics*, 68:1–18.

Kneebone, R. and White, K. (2015). An overview of social assistance trends in canada. *Welfare reform in Canada: Provincial social assistance in comparative perspective*, pages 53–92.

Kostøl, A., Mogstad, M., Setzler, B., et al. (2019). Disability benefits, consumption insurance, and household labor supply. *American Economic Review*, 109(7):2613–54.

Liebman, J. B. (2015). Understanding the increase in disability insurance benefit receipt in the united states. *Journal of Economic Perspectives*, 29(2):123–150.

Loprest, P. and Maag, E. (2007). The relationship between early disability onset and education and employment. *Journal of Vocational Rehabilitation*, 26(1):49–62.

Low, H., Meghir, C., and Pistaferri, L. (2010). Wage risk and employment risk over the life cycle. *American Economic Review*, 100(4):1432–67.

Low, H. and Pistaferri, L. (2015). Disability insurance and the dynamics of the incentive insurance trade-off. *American Economic Review*, 105(10):2986–3029.

Lundborg, P., Nilsson, A., and Rooth, D.-O. (2014). Adolescent health and adult labor market outcomes. *Journal of Health Economics*, 37:25–40.

Maytree (2018). Social assistance summaries 2018.

Meyer, B. D., Mok, W. K., and Sullivan, J. X. (2009). The under-reporting of transfers in household surveys: its nature and consequences. Technical report, National Bureau of Economic Research.

Michaud, A. and Wiczer, D. (2018). The disability option: Labor market dynamics with macroeconomic and health risks. Technical report, CHCP Working Paper.

Millard, B. (2023). The longitudinal effects of disability types on incomes and employment. In *2023 Stata Conference*, number 07. Stata Users Group.

Milligan, K. (2016). Canadian tax and credit simulator. Database, software and documentation, Version 2016-2.

Milligan, K. and Schirle, T. (2019). Push and pull: Disability insurance, regional labor markets, and benefit generosity in canada and the united states. *Journal of Labor Economics*, 37(S2):S289–S323.

Mori, H. (2016). Essays on human capital complementarities.

Morris, S. P., Fawcett, G., Brisebois, L., and Hughes, J. (2018). A demographic, employment and income profile of canadians with disabilities aged 15 years and over, 2017.

Navarro, S. and Zhou, J. (2017). Identifying agent's information sets: An application to a lifecycle model of schooling, consumption and labor supply. *Review of Economic Dynamics*, 25:58–92.

Office of the Auditor General of Canada (2015). Report 6-canada pension plan disability program. `https://www.oag-bvg.gc.ca/internet/english/parl_oag_201602_06_e_41063.html`.

Prinz, D., Chernew, M., Cutler, D., and Frakt, A. (2018). Health and economic activity over the lifecycle: Literature review.

Shorrocks, A. F. et al. (2013). Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *Journal of Economic Inequality*, 11(1):99–126.

Stern, S. (1989). Measuring the effect of disability on labor force participation. *Journal of human Resources*, pages 361–395.

Torjman, S. and Makhoul, A. (2016). *Disability supports and employment policy*. Caledon Institute of Social Policy.

Appendix

A.1. Model Parameters

Table 9: Summary of Model Parameters.

Parameter	Description
Individual Heterogeneity	
$\bar{a}^{d_0}$	Mean of endowed ability distribution
$\sigma^2_{a^{d_0}}$	Variance of endowed ability distribution
$\gamma^{d_0,t}_{i,j}$	Disability risk for transitioning from $d_t = i, d_{t-1} = j$
Earnings Process	
ϕ	Direct effect of disability on earnings
$\mu^{d_0,s}_1$	Return to potential experience for d_0
$\mu^{d_0,s}_2$	Return to squared potential experience for d_0
$\sigma^2_{\xi^{s,d_0}}$	Variance of productivity shock for (s, d_0)
$\bar{\xi}_0$	Mean of initial productivity shock
$\sigma^2_{\xi_0}$	Variance of initial productivity shock
$h^s_{d_0}$	Return to university for d_0
Utility Parameters	
κ	Coefficient of relative risk aversion
θ	Utility cost of disability
η	Utility cost of working
β	Discount factor
Policy Parameters	
π^s	Probability of DI acceptance
π^{SA}	Probability of receiving additional disability benefits in SA
$\pi^{d_0,s}_{ret}$	Retirement risk for s, d_0 at age t
C^{s,d_0}_{App}	Utility cost of DI application
Labour Market Environment	
$\delta^{d,s}$	Exogenous Job Destruction
$\lambda^{s,d}$	Exogenous Job Arrival Rate
ρ	Disability reassessment rate
Psychic Cost of School	
$\sigma^2_{\epsilon_\psi}$	Variance of idiosyncratic psychic cost of university
g_0	Mean of psychic cost of schooling
g_1	Difference in mean of psychic cost of schooling for $d_0 = 1$

A.2.1 Details on Numerical Solution to the Model

For a given set/guess of the model's parameters, the model is solved starting with decisions in the terminal period, i.e., retirement, and then iterating backwards to solve individual's decisions at each period, conditional on state variables. The value of the terminal period is deterministic conditional on the state variables. Moving back to T-1,

1. For each realized combination of discrete state variables (time varying and fixed), the continuation value (EMAX) is calculated on a discrete grid of the continuous state variables. Continuous state variables are $(a_i, \epsilon_{it}, e_{i,t-1})$, which can be reduced to $(W_{it}, e_{i,t-1})$ given a and ϵ only affect earnings growth.

2. The continuation value depends on expectations over both discrete and continuous random variables. Given the assumed normal distribution of the productivity shock, ϵ_{it+1}, I numerically integrate using Gauss-Hermite quadrature.

3. The continuation value between a discretized grid of continuous state variables is interpolated using a bilinear interpolation algorithm. This procedure is repeated moving backwards to T-2, T-3, etc. where the interpolated conditional EMAX is used in place of the continuation value.

4. This process yields 32 $\times$ 47 distinct interpolation functions, which represent the EMAX conditional on a given combination of discrete state variables for each of the 47 periods.

A.3. Sample Survey Questions on Limitations to Daily Activities

Table 10: Questions used to Measure Limitations to Daily Activities

Questions to Derive Aggregate Physical Disability
How much difficulty do you have walking on a flat surface for 15 minutes without resting?
How much difficulty do you have walking up or down a flight of stairs, about 12 steps without resting?
How much difficulty do you have reaching in any direction, for example, above your head?
How much difficulty do you have using your fingers to grasp small objects like a pencil or scissors?
Do you have pain that is always present?
Questions to Derive Mental-Cognitive Disability
Do you think you have a condition that makes it difficult in general for you to learn? This may include learning disabilities such as dyslexia, hyperactivity, attention problems, etc..
Has a teacher, doctor or other health care professional ever said that you had a learning disability?
Has a doctor, psychologist or other health care professional ever said that you had a developmental disability or disorder? This may include Down syndrome, autism, Asperger syndrome, mental impairment due to lack of oxygen at birth, etc..
Do you have any ongoing memory problems or periods of confusion? Please exclude occasional forgetfulness such as not remembering where you put your keys.
Do you have any emotional, psychological or mental health conditions? These may include anxiety, depression, bipolar disorder, substance abuse, anorexia, etc..

Source: Table comes directly from Grondin, C. (2016). A new survey measure of disability: The Disability Screening Questions (DSQ). Statistics Canada.

A.4. Tax and Transfer system

Parameters for the income tax brackets and marginal tax rates were derived from the Canadian Tax and Transfer Simulator (Milligan, 2016). I cap the upper threshold to tax brackets to give me 5 distinct tax brackets. The weights are based on the joint density of calendar year and province in my sample. The income tax regime in my model is shown in Table 11.

Table 11: Tax Brackets and Marginal Tax Rates.

Income Bracket	Tax Rate
[0, 30805]	0.2280
[30805, 46586]	0.2944
[46586, 64178]	0.3433
[64178, 68066]	0.3621
[68066, ∞]	0.3833

A.5. Auxiliary Moments

Tables 12 to 26 display the full set of auxiliary moments used in estimation. Each table reports the moments calculated in the data, the moments calculated using data simulated with the model, the difference in the data moment and the simulated moments, and the standard error in from the data. Estimation consists of 201 moments. Almost all moments are separated by early disability status and education level.

Table 12: Post-Secondary Distribution

	Data	Simulation	Standard Error	Diff.
$Frac(s = 1 \mid d_0 = 1)$	0.6380	0.6372	0.0120	0.0009
$Frac(s = 1 \mid d_0 = 0)$	0.4669	0.4667	0.0370	0.0002

Table 13: Education Regressions

	Data	Simulation	Standard Error	Diff.
d_0	-0.1000	-0.0885	0.0395	0.0114
$\hat{v}$	0.1799	0.1366	0.0185	0.0433
Intercept	-1.0250	-0.6457	0.1745	0.3793
σ^2_ψ	0.2161	0.2269	0.0040	0.0108

Table 14: Employment Rates

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
$Fr(L_{it} = 1 \mid d^*_{it} = 0, t < 45)$	0.8743	0.8456	0.0030	0.0287
$Fr(L_{it} = 1 \mid d^*_{it} = 0, t \geq 45)$	0.7974	0.8283	0.0040	0.0309
$Fr(L_{it} = 1 \mid d^*_{it} = 1, t < 45)$	0.6703	0.9098	0.0150	0.2395
$Fr(L_{it} = 1 \mid d^*_{it} = 1, t \geq 45)$	0.4786	0.6988	0.0090	0.2202
		Not Early, Post-Secondary		
$Fr(L_{it} = 1 \mid d^*_{it} = 0, t < 45)$	0.9076	0.9292	0.0020	0.0216
$Fr(L_{it} = 1 \mid d^*_{it} = 0, t \geq 45)$	0.8504	0.9540	0.0030	0.1037
$Fr(L_{it} = 1 \mid d^*_{it} = 1, t < 45)$	0.8307	0.9628	0.0080	0.1322
$Fr(L_{it} = 1 \mid d^*_{it} = 1, t \geq 45)$	0.6385	0.8913	0.0070	0.2528
		Early-onset, Low Education		
$Fr(L_{it} = 1 \mid t < 45)$	0.5213	0.6269	0.0140	0.1056
$Fr(L_{it} = 1 \mid t \geq 45)$	0.4799	0.5954	0.0200	0.1155
		Early-onset, Post-Secondary		
$Fr(L_{it} = 1 \mid t < 45)$	0.8152	0.7702	0.0090	0.0450
$Fr(L_{it} = 1 \mid t \geq 45)$	0.6107	0.6843	0.0180	0.0736

Table 15: Initial Employment

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
$Fr(L_{it} = 1 \mid t \leq 3)$	0.8095	0.6222	0.0090	0.1873
		Not Early, Post-Secondary		
$Fr(L_{it} = 1 \mid 4 \geq t \leq 6)$	0.8624	0.7896	0.0060	0.0728
		Early-onset, Low Education		
$Fr(L_{it} = 1 \mid t \leq 3)$	0.5794	0.4410	0.0280	0.1383
		Early-onset, Post-Secondary		
$Fr(L_{it} = 1 \mid 4 \geq t \leq 6)$	0.8152	0.4388	0.0220	0.3763

Table 16: Flows Into and Out of Employment

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
$Fr(L_{it} = 0 \| L_{it-1} = 1, t < 45)$	0.0464	0.0325	0.0020	0.0139
$Fr(L_{it} = 0 \| L_{it-1} = 1, t \geq 45)$	0.0402	0.0507	0.0020	0.0104
$Fr(L_{it} = 1 \| L_{it-1} = 0, t < 45)$	0.0503	0.0489	0.0020	0.0015
$Fr(L_{it} = 1 \| L_{it-1} = 0, t \geq 45)$	0.0280	0.0292	0.0020	0.0011
		Not Early, Post-Secondary		
$Fr(L_{it} = 0 \| L_{it-1} = 1, t < 45)$	0.0373	0.0197	0.0010	0.0176
$Fr(L_{it} = 0 \| L_{it-1} = 1, t \geq 45)$	0.0392	0.0268	0.0010	0.0124
$Fr(L_{it} = 1 \| L_{it-1} = 0, t < 45)$	0.0468	0.0622	0.0010	0.0154
$Fr(L_{it} = 1 \| L_{it-1} = 0, t \geq 45)$	0.0245	0.0207	0.0010	0.0038
		Early-onset, Low Education		
$Fr(L_{it} = 0 \| L_{it-1} = 1, t < 45)$	0.0799	0.0628	0.0070	0.0172
$Fr(L_{it} = 0 \| L_{it-1} = 1, t \geq 45)$	0.0286	0.0700	0.0060	0.0413
$Fr(L_{it} = 1 \| L_{it-1} = 0, t < 45)$	0.0670	0.0749	0.0070	0.0079
$Fr(L_{it} = 1 \| L_{it-1} = 0, t \geq 45)$	0.0192	0.0565	0.0050	0.0373
		Early-onset, Post-Secondary		
$Fr(L_{it} = 0 \| L_{it-1} = 1, t < 45)$	0.0504	0.0331	0.0050	0.0173
$Fr(L_{it} = 0 \| L_{it-1} = 1, t \geq 45)$	0.0454	0.0551	0.0080	0.0097
$Fr(L_{it} = 1 \| L_{it-1} = 0, t < 45)$	0.0582	0.0718	0.0050	0.0136
$Fr(L_{it} = 1 \| L_{it-1} = 0, t \geq 45)$	0.0267	0.0350	0.0050	0.0083

Table 17: DI Rates, Flows, and Coverage

	Data	Simulation	Standard Error	Diff.
DI Coverage				
$Fr(DI_{it} = 1 \| d_{it}^{*} = 1, d_{i0} = 0, s = 0)$	0.0238	0.0250	0.0010	0.0013
$Fr(DI_{it} = 1 \| d_{it}^{*} = 1, d_{i0} = 0, s = 1)$	0.0085	0.0013	0.0004	0.0072
$Fr(DI_{it} = 1 \| d_{it}^{*} = 1, d_{i0} = 1, s = 0)$	0.0396	0.0313	0.0050	0.0083
$Fr(DI_{it} = 1 \| d_{it}^{*} = 1, d_{i0} = 1, s = 1)$	0.0407	0.0350	0.0040	0.0057
DI Flows				
$Fr(D_{it} = 1 \| DI_{it-2} = 0, d_{it} = 0, s = 0)$	0.0035	0.0028	0.0004	0.0007
$Fr(D_{it} = 1 \| DI_{it-2} = 0, d_{it} = 0, s = 1)$	0.0015	0.0002	0.0002	0.0013
$Fr(D_{it} = 1 \| DI_{it-2} = 0, d_{it} = 1, s = 0)$	0.0042	0.0039	0.0010	0.0003
$Fr(D_{it} = 1 \| DI_{it-2} = 0, d_{it} = 1, s = 1)$	0.0051	0.0047	0.0010	0.0004
DI Composition				
$Fr(d_{it} = 0, s = 1 \| DI_{it} = 1)$	0.3715	0.0189	0.0130	0.3525
$Fr(d_{it} = 1, s = 0 \| DI_{it} = 1)$	0.0686	0.3974	0.0080	0.3289
$Fr(d_{it} = 1, s = 1 \| DI_{it} = 1)$	0.0847	0.3638	0.0080	0.2791

Table 18: Coefficients from Linear Regression for DI rate

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
age	0.0068	0.0082	0.0024	0.0014
age^2	-0.0002	-0.0003	0.0001	0.0001
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	-0.0630	-0.0733	0.0267	0.0102
		Not Early, Post-Secondary		
age	0.0062	0.0017	0.0015	0.0045
age^2	-0.0002	-0.0001	0.0000	0.0001
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	-0.0682	-0.0180	0.0184	0.0502
		Early-onset, Low Education		
age	0.0211	0.0108	0.0141	0.0103
age^2	-0.0007	-0.0004	0.0004	0.0004
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	-0.1933	-0.1014	0.1534	0.0918
		Early-onset, Post-Secondary		
age	-0.0032	0.0271	0.0140	0.0303
age^2	0.0001	-0.0008	0.0004	0.0009
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	0.0385	-0.2759	0.1672	0.3144

Table 19: Coefficients from Linear Regression for DI flow

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
age	-0.0023	-0.0032	0.0008	0.0009
age^2	0.0001	0.0001	0.0000	0.0000
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	0.0270	0.0375	0.0096	0.0105
		Not Early, Post-Secondary		
age	0.0006	-0.0001	0.0006	0.0007
age^2	0.0000	0.0000	0.0000	0.0000
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	-0.0068	0.0012	0.0075	0.0080
		Early-onset, Low Education		
age	-0.0040	-0.0035	0.0038	0.0005
age^2	0.0001	0.0001	0.0001	0.0000
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	0.0452	0.0420	0.0419	0.0031
		Early-onset, Post-Secondary		
age	0.0048	-0.0036	0.0058	0.0084
age^2	-0.0001	0.0001	0.0002	0.0002
age^3	0.0000	0.0000	0.0000	0.0000
Intercept	-0.0577	0.0443	0.0695	0.1020

Table 20: Mean and Variance of Initial Earnings (conditional on working)

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
$E(LnW \mid t \leq 3)$	9.3900	9.6361	0.0180	0.2461
$Var(LnW \mid t \leq 3)$	9.7000	9.8450	0.0120	0.1450
		Not Early, Post-Secondary		
$E(LnW \mid 4 \leq t \leq 6)$	9.2000	9.2167	0.0520	0.0167
$Var(LnW \mid 4 \leq t \leq 6)$	9.5300	9.6928	0.0470	0.1628
		Early-onset, Low Education		
$E(LnW \mid 1 \leq t \leq 3)$	0.4857	0.2244	0.0150	0.2613
$Var(LnW \mid 1 \leq t \leq 3)$	0.5257	0.2788	0.0100	0.2469
		Early-onset, Post-Secondary		
$E(LnW \mid 4 \leq t \leq 6)$	0.5322	0.4206	0.0360	0.1116
$Var(LnW \mid 4 \leq t \leq 6)$	0.6182	0.2738	0.0390	0.3444

Table 21: Fixed Effect Earnings Regression

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
$d_{it}*$	-0.0271	-0.0384	0.0400	0.0113
PE	0.1099	0.1042	0.0062	0.0057
PE^2/100	-0.2032	-0.2209	0.0120	0.0177
$E(v)$	9.0371	9.2338	0.0170	0.1967
$Var(v)$	0.3442	0.3714	0.0140	0.0272
$V(\xi)$	0.1978	0.1705	0.0050	0.0273
		Not Early, Post-Secondary		
$d_{it}*$	-0.0241	-0.0359	0.0236	0.0118
PE	0.1449	0.1237	0.0041	0.0212
PE^2/100	-0.2852	-0.2557	0.0096	0.0295
$E(v)$	9.2047	9.4936	0.0130	0.2889
$Var(v)$	0.3643	0.3448	0.0140	0.0195
$V(\xi)$	0.1796	0.1625	0.0030	0.0171
		Early-onset, Low Education		
PE	0.1043	0.0996	0.0267	0.0047
PE^2/100	-0.1971	-0.1730	0.0637	0.0241
$E(v)$	8.7614	8.7411	0.0610	0.0203
$Var(v)$	0.4713	0.4561	0.0530	0.0152
$V(\xi)$	0.2723	0.1783	0.0210	0.0940
		Early-onset, Post-Secondary		
PE	0.1330	0.1218	0.0185	0.0112
PE^2/100	-0.3227	-0.2564	0.0361	0.0662
$E(v)$	9.2447	8.9793	0.0530	0.2653
$Var(v)$	0.3875	0.5359	0.0440	0.1484
$V(\xi)$	0.2197	0.1736	0.0190	0.0461

Table 22: Fixed Effect Quantiles

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
Q10	8.2280	8.4570	0.0360	0.2290
Q25	8.6730	8.7953	0.0280	0.1223
Q50	9.0940	9.2053	0.0220	0.1113
Q75	9.4320	9.6526	0.0200	0.2206
Q90	9.7460	10.0568	0.0230	0.3108
		Not Early, Post-Secondary		
Q10	8.4320	8.7254	0.0260	0.2934
Q25	8.8550	9.0872	0.0200	0.2322
Q50	9.2820	9.5064	0.0140	0.2244
Q75	9.6200	9.9034	0.0160	0.2834
Q90	9.8970	10.2641	0.0130	0.3671
		Early-onset, Low Education		
Q10	7.9110	7.9041	0.1250	0.0069
Q25	8.2830	8.2409	0.0790	0.0421
Q50	8.7710	8.6839	0.0790	0.0871
Q75	9.1860	9.1945	0.1060	0.0085
Q90	9.6820	9.6920	0.0770	0.0100
		Early-onset, Post-Secondary		
Q10	8.4280	8.0509	0.0870	0.3771
Q25	8.8630	8.4465	0.0880	0.4165
Q50	9.2930	8.9542	0.0800	0.3388
Q75	9.6770	9.4824	0.0540	0.1946
Q90	10.0950	9.9502	0.1080	0.1448

Table 23: Pooled Earnings Regression

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
age	0.1221	0.1231	0.0027	0.0010
$age^2/100$	-0.1303	-0.1417	0.0033	0.0115
intercept	7.6334	7.7532	0.0509	0.1198
$Var(\epsilon)$	0.4749	0.4461	0.0060	0.0288
		Not Early, Post-Secondary		
age	0.1729	0.1978	0.0023	0.0249
$age^2/100$	-0.1843	-0.2181	0.0028	0.0338
intercept	6.8515	6.4424	0.0444	0.4090
$Var(\epsilon)$	0.4868	0.4274	0.0040	0.0595
		Early-onset, Low Education		
age	0.0957	0.0941	0.0112	0.0016
$age^2/100$	-0.0890	-0.0931	0.0145	0.0041
intercept	7.7130	7.8965	0.1972	0.1835
$Var(\epsilon)$	0.5871	0.5333	0.0240	0.0539
		Early-onset, Post-Secondary		
age	0.1900	0.1150	0.0117	0.0750
$age^2/100$	-0.2040	-0.1198	0.0145	0.0842
intercept	6.3307	7.6730	0.2214	1.3423
$Var(\epsilon)$	0.5448	0.5503	0.0230	0.0055

Table 24: Earnings Quantiles

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
Q10	9.1485	9.3083	0.0140	0.1598
Q25	9.8574	9.6725	0.0090	0.1850
Q50	10.3890	10.1515	0.0050	0.2375
Q75	10.7706	10.6752	0.0040	0.0954
Q90	11.0604	11.1618	0.0040	0.1014
Average	32300	34728	100	2428
		Not Early, Post-Secondary		
Q10	9.5252	9.6320	0.0110	0.1068
Q25	10.2400	10.0702	0.0060	0.1698
Q50	10.7515	10.5836	0.0030	0.1679
Q75	11.1258	11.1080	0.0030	0.0178
Q90	11.4164	11.5610	0.0040	0.1446
Average	50900	51529	100	629
		Early-onset, Low Education		
Q10	8.5564	9.1326	0.0720	0.5762
Q25	9.2780	9.4984	0.0440	0.2204
Q50	9.9711	9.9867	0.0300	0.0155
Q75	10.4545	10.5723	0.0220	0.1178
Q90	10.9187	11.1405	0.0320	0.2218
Average	26000	32095	500	6095
		Early-onset, Post-Secondary		
Q10	9.1270	9.2898	0.0620	0.1628
Q25	9.8679	9.6657	0.0330	0.2021
Q50	10.4940	10.1807	0.0230	0.3133
Q75	10.9096	10.7706	0.0170	0.1390
Q90	11.2424	11.3123	0.0180	0.0699
Average	40400	38505	600	1895

Table 25: Variance and Autocorrelation of residuals from lagged earnings regression

	Data	Simulation	Standard Error	Diff.
		Not Early, Low Education		
$Var(\epsilon_t^L)$	0.1709	0.1553	0.0040	0.0155
$Corr(\epsilon_t^L, \epsilon_{t-1}^L)$	-0.0904	-0.0002	0.0120	0.0902
		Not Early, Post-Secondary		
$Var(\epsilon_t^L)$	0.1572	0.1478	0.0030	0.0095
$Corr(\epsilon_t^L, \epsilon_{t-1}^L)$	-0.0712	-0.0474	0.0080	0.0238
		Early-onset, Low Education		
$Var(\epsilon_t^L)$	0.2354	0.1636	0.0180	0.0718
$Corr(\epsilon_t^L, \epsilon_{t-1}^L)$	-0.1217	-0.0002	0.0480	0.1215
		Early-onset, Post-Secondary		
$Var(\epsilon_t^L)$	0.1808	0.1603	0.0140	0.0205
$Corr(\epsilon_t^L, \epsilon_{t-1}^L)$	-0.0459	-0.0493	0.0530	0.0034

Table 26: Pre-DI Earnings Quantiles and Rate Not Working

	Data	Simulation	Standard Error	Diff.
		Low Education		
Q15	9.0825	8.9565	0.4070	0.1261
Q25	9.5178	9.1592	0.2230	0.3586
Q50	10.1849	9.4158	0.1450	0.7691
Q75	10.5713	9.6450	0.1200	0.9263
Q90	10.8780	9.8441	0.1450	1.0339
$(1 - Fr(L_{it} = 1 \vert DI_{it+2} = 1))$	0.3785	0.6027	0.0520	0.2243
		Post-Secondary		
Q15	9.2496	8.9683	0.2360	0.2813
Q25	9.4572	9.1705	0.1710	0.2867
Q50	10.4073	9.3953	0.1830	1.0120
Q75	10.8396	9.6551	0.0840	1.1845
Q90	11.1110	9.8221	0.1760	1.2889
$(1 - Fr(L_{it} = 1 \vert DI_{it+2} = 1))$	0.3350	0.5706	0.0470	0.2355
		Early-onset		
Q10	8.8099	8.8750	0.2930	0.0651
Q25	9.5178	9.0857	0.1580	0.4321
Q50	10.2471	9.4374	0.0850	0.8097
Q75	10.7515	9.6684	0.0690	1.0831
Q90	11.0867	9.7766	0.1110	1.3101
$(1 - Fr(L_{it} = 1 \vert DI_{it+2} = 1))$	0.3280	0.6273	0.0360	0.2993

A.6. Additional Decomposition: Ceteris Paribus decomposition

This section analyzes the drivers of the estimated education gap, hereby referred to as the baseline gap. To hone in on the contribution of different parameters, I sequentially shut down model features or equate parameters that differ by d_0, by setting them to the value estimated for not early disabled individuals while keeping the others at their estimated values. I then resolve the model under this alternate environment and analyze how simulated individuals change their behaviour, relative to the baseline. It is important to note that this strategy captures the average effect of parameter differences, which includes complementary interactions with the other model parameters. The results from this decomposition exercise are reported in Table 27.

Table 27: Decomposing the Simulated Gap in Post-Secondary Education by d_0

	Fraction in Post-Secondary				
	Not Early Disabled	Early-Onset	Gap	Net Baseline Gap	% of Baseline Gap
1. Baseline Gap	0.637	0.467	0.170		
(Non-policy) Counterfactual Scenario					
2. Disability Risk	0.638	0.471	0.167	0.003	2.019
3. Psychic Cost	0.637	0.474	0.163	0.007	4.139
4. Ability Endowment	0.637	0.490	0.147	0.024	13.84
5. Return to school	0.637	0.545	0.092	0.078	45.77
6. Productivity Shocks	0.637	0.469	0.169	0.002	1.064
7. Direct effect on Earnings	0.638	0.468	0.170	0.001	0.358
8. Labour Market Risks	0.637	0.456	0.181	-0.010	-6.150
9. Utility Cost of Disability	0.630	0.490	0.140	0.031	17.91
(Policy) Counterfactual Scenario					
10. SA-D	0.646	0.502	0.144	0.026	15.574
11. DI	0.639	0.467	0.172	0.009	-1.047
12. SA-D and DI	0.648	0.503	0.145	-0.005	14.996