Early-Onset Disability, Education Investments, and Social Insurance

By Robert Millard*

The University of Western Ontario

April 18, 2022

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. In Canada, this equates to an eighteen percentage point gap in post-secondary attainment between these populations. This gap is likely related to how an early-onset disability affects the cost and return to investing in education and the availability of additional income through social insurance policies. In this paper, 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 that the effect of an early-onset disability on the financial return to education plays the most prominent role in driving the education differential between the early-onset and non-disabled individuals. I also find approximately 15% of the gap in educational attainment is due to additional resources available via SA for individuals with a disability. DI is more relevant for older ages and has trace effects on education. Through counterfactual experiments, I find education investments and employment rates are inversely related to the generosity of SA, and there is a tradeoff between individual welfare and the moral hazard from the generosity of SA. 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.

*Ph.D. Candidate, The University of Western Ontario. 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 for many helpful discussions. I am also grateful for helpful comments from participants of various seminars and reading groups in Economics department at the University of Western Ontario, notably Lance Lochner, Baxter Robinson, and Sergio Ocampo Diaz. In addition, thank you for discussions from various speakers visiting Western over the previous few years

1 Introduction

The onset of a work-limiting disability has large and persistent consequences to individual economic independence and overall welfare.[1] Disability limits the set of productive tasks one can engage in to make a living and is a primary source of labour market inequality among working-age adults.[2] The individual effects of disability culminate into a substantial cost on public infrastructure, which bears a large financial burden of providing support to those affected by disability.[3] The barriers, costs, and limitations from disability make this group's behaviour especially sensitive to social welfare policies, which offer financial transfers to help insure against disability shocks. The economic costs of disability can be exacerbated for conditions that onset before age eighteen, hereby referred to as early-onset.

In this paper, I investigate the factors driving education choices for early-onset individuals to gain insight into the link between childhood health and adult outcomes. Bad health early in life generally results in lower attainment of post-secondary education.[4] This paper focuses on Canada, where early-onset individuals, who represent about a quarter of the Canadian population with a disability, are 18 percentage points less likely to earn a post-secondary certificate relative to their non-disabled counterparts. Early-onset disabilities are present at critical periods of skill development and investment in human capital during schooling. The consequences of disrupting human capital accumulation early in life extend to adult labour market outcomes, as education is a crucial determinant of financial independence and labour market success.[5] Better understanding the relationship between early-onset disability and education investments is important to improve outcomes and welfare for this demographic in the labour market.

Education is an important mechanism for early-onset individuals to enhance their labour market productivity and offset or overcome the barriers induced by their disability. Education choices depend on one's expectation of future employability, income, and labour market risks.[6] Individuals with lower levels of

[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, Meyer and Mok 2012, Millard 2021).

[2] Disability rates have been rising over the past few decades in Canada, as well as most developed countries. 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 evolution of the definition of what constitutes disability and changes in individual reporting behaviour. For more details on the economic position of Canadians with disabilities see Morris et al. (2017) and Cossette and Duclos (2001).

[3] A review of the growing public burden of the US disability system can be found in Autor and Duggan (2006).

[4] 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.

[5] Disabling conditions early in life have been found to stunt earnings growth, be associated with low labour force attachment and greater dependence on transfers. Some articles researching or surveying research on the impact of early health on adult outcomes include Case and Paxton (2010), Almond and Currie (2010, 2018), Lundbourg (2014), and Prinz et al. (2018).

[6] Education choice is often modelled as a discrete choice over expected values associated with each education level. The expectations can be functions of labour market conditions, uncertainty, and policy. Individuals choose the education level that gives them the greatest expected discounted lifetime value (Cunha, Heckman and Navarro 2006, Keane, Todd and Wolpin 2011).

completed education have a greater likelihood of receiving low earnings and experiencing unemployment in working life. Social insurance (SI) provides resources to ensure individual consumption does not substantially decline during these difficult labour market states. However, SI may also diminish the incentive to invest in one's education by raising the relative value of low earning states. Measurements of the size of this disincentive have substantial policy relevance. Individuals affected by disability have high rates and long spells of attachment to government transfers (Torjman 2017). Separating the role of SI from other factors that influence education investments is essential to design policy to improve education outcomes, promote investment into individual productivity, and enhance economic independence for individuals affected by early-onset disabilities.

This paper analyzes the relative importance of SI policy, among other factors, in affecting post-secondary investments of early-onset individuals. I build and estimate a structural life-cycle model of education, employment, and SI receipt. Education is a costly investment into productivity. Rational, forward looking individuals base their decisions on the expected costs and labour market return to each education level. There are many ways that an early-onset disability can affect the return to education. The efficiency of human capital accumulation during post-secondary education and in the labour market may be limited by an early-onset disability, lowering the financial return to post-secondary education. Additionally, labour market frictions, which increase the likelihood of unemployment, vary by education level and disability status, further affecting the returns to post-secondary investment. SI benefits raise the value of not working and alter the incentive to invest in education differently for individuals with a disability. Hence, early-onset individuals in the model may put more weight on SI in comparing the expected values of education levels.

A structural model allows me to specify and separately analyze a comprehensive set of factors that may lower the return to schooling for early-onset individuals. Life-cycle information on SI applications in conjunction with education and labour market behaviour is scarce. Furthermore, the decision to apply for SI, which is central to this research question, is unobserved. The structural approach specifies a model that represents endogenous or latent choices and their relationship to the models parameters and individual's state variables. For instance, disability is a determinant of education, labour supply, and SI receipt. The model specifies how disability affects these, and therefore, how these decisions relate to each other via disability status. To measure the importance of features, I can manipulate parameters that differ by early disability status and use the estimated model to predict changes in behaviour. Moreover, I can manipulate policy parameters to analyze the effects of counterfactual reforms to the policy environment on individual's behaviour and welfare.

The model is representative of the Canadian labour market and SI policy environment. The programs making up the Canadian SI system for individuals affected by disability operates as a set of separate welfare initiatives rather than a unified system of integrated policies (Torjman 2017). This feature of the Canadian environment is particularity useful for evaluating policy reforms, as the effect on behaviour is less confounded by interactions with other programs.[7]

In Canada, disability status may grant access to additional SI resources for working-aged adults. I focus on two of the main income assistance programs available for individuals with disabilities: disability insurance (DI) and social assistance (SA), the latter of which allocates additional resources for individuals with a disability (SA-D). These programs are designed to aid with the disadvantage and increased uncertainty associated with a disability. Individuals with disabilities are at a higher risk of requiring assistance and have a larger set of resources available. Hence, the disincentive effects of SI policy on the return to education may be amplified for this demographic. Additionally, some programs, notably DI, condition on education level in determining the program's eligibility, which can add to disincentives (Government of Canada 2018b). Accounting for the incentives created by SI is essential to understand the factors driving the educational attainment of individuals with an early-onset disability.

I estimate my 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 demographics. LISA is linked to a 36-year panel of administrative tax records containing disaggregated measures of personal incomes, taxes, and transfers. These survey data allow me to separately analyze the behavioural incentives of SA and DI. Additionally, the use of 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 2011, Meyer, Mok and Sullivan 2009). The merged survey and administrative information facilitate the creation of a rich dataset with detailed health measures, demographics, and incomes. The combination of these data make LISA well suited to study the life-cycle behaviour of early-onset individuals. Furthermore, LISA includes a large subsample of individuals with early-onset disabilities, which previously was a limitation for researching this group.[8]

I measure disability based on limitations to activities of daily living (LADLs), rather than specific health

[7]For instance, the main disability insurance program in the US, SSDI, may grant access to medical insurance, which can affect the incentive to pursue SSDI. Whereas Canadians are covered by universal healthcare and applications to the various disability programs are separately administered.

[8]Studying early health conditions has historically been difficult as it requires data on early life conditions in addition to information on the outcomes of interest, which are often from adulthood. LISA provides me with information on Canadians over much of their life-cycle. LISA's detailed information makes it an ideal dataset to study the importance of Canada's public policies and labour market conditions on the education decision of people with early-onset disabilities.

conditions. LADLs capture a key intermediate step in the mapping from health to productivity. It may not always be evident if, or how, a given health condition, such as diabetes, affects observed behaviour. Instead, measures of how one's diabetes affects their ability to perform daily activities give more clarity into the behaviour this condition restricts. Additionally, this stance on disability is useful for studying education, which is a direct investment into productive ability.

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 actual data, and I fit a positive average earnings difference by education level. The estimated parameters find the not early disabled population has a higher average ability endowment and a larger financial return to education. I fit the rate of DI recipt and recover estimated DI acceptance probabilities that are consistent with the probability of acceptance observed in a 2015 audit of the Canadian Federal DI program (Office of the Auditor General of Canada 2015).

I use the estimated model to decompose the education gap between early-onset individuals and their not early disabled counterparts. I find the largest contributor is differences in the financial returns to education, accounting for 43% of the observed gap. SI policy plays an important role in driving the education gap because of the added generosity of SA-D benefits. SA-D benefits account for 15.6% of the education gap. DI is more relevant at older ages and has minor effects on education. DI policy helps to offset the education gap, because individuals that are not early disabled are relatively more sensitive to DI incentives than early-onset individuals. Individuals who are not early disabled have higher earnings on average, which increases the value of DI. Furthermore, I recover some evidence that the DI acceptance probability depends on education level, but this does not affect schooling choices. Ability and preferences also play important roles, accounting for 13% and 17% of the gap, respectively. Finally, labour market frictions increase the likelihood of post-secondary for early-onset individuals. This is because the gain to the likelihood of being employed, through job arrival and destruction rates, from completing post-secondary is much larger for early-onset individuals.

I then evaluate the implications on education and life-cycle behaviour for a set of counterfactual policy reforms. On the one hand, SA has been criticized for not providing sufficient resources to cover the cost of living for it's recipients (Tweddle and Aldridge 2017, Hillel 2020) On the other hand, critiques of SA are with the disincentives created by the program (Béland 2015). Increasing the generosity of SA increases the moral hazard of the program, resulting in less post-secondary attainment and lower rates of employment in the labour market. This reform results in higher rates of SA and DI because individuals flow to DI from

SA. Increasing the generosity of DI does not affect schooling but has minor effects on employment. The decomposition exercise found SA to have more of a disincentive on education than DI. Reallocating resources between programs may reduce the disincentive effects while retaining the overall insurance of the SI environment. Reallocating resources from SA to DI results in higher rates of post-secondary and employment, but at the expense of individual welfare.

The final policy I consider is a grant for early-onset individuals during their post-secondary schooling. This counterfactual policy directly incentivizes investment in productivity, resulting in a 6.87 percentage point increase in post-secondary enrolment. Early-onset individuals in this scenario have higher average earnings and employment rates. Furthermore, this policy helps pay for itself through lower rates of SA and greater tax revenues from the higher employment and average earnings. These results reinforce the importance of education in helping equip individuals to thrive in the workforce and be less reliant on SI.

A broader understanding of the role of disability, uncertainty, and SI policy on education investments and labour market outcomes is of considerable interest. I frame my research contribution in three 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 their entire life. I focus on education choice as a mechanism by which individuals affected by a disability early in life can improve their labour market outcomes in adulthood.[10] I emphasize the role of labour market policy in incentivizing higher education investments.

Second, I contribute to a body of literature interested in how individual behaviour is affected by the parameters of SI policy. This paper fits among a number of studies that structurally model how disability policy drives labour market choices.[11] To my knowledge, this is the first study to measure an insurance-incentive trade-off of SA and DI with respect to educational investments. My results offer important insight into thinking about the design of DI and welfare programs when considering people affected by an early-onset disability.[12]

Third, I contribute to the literature on the relationship between human capital investments, labour market conditions, and SI policies. Again, this paper aligns with various structural models linking education rates to the labour market environment.[13] The idea is that risks and public policies create incentives that

[9] For examples and surveys of related studies see Case et al (2010), Almond and Currie (2010,2018), Lundbourg (2014), Mori (2016), Prinz et al (2018), Millard (2021).

[10] Mori (2016) conducts a similar structural analysis of how an early-onset disability affects education investments. My study differs in that I account for the role of SI policy.

[11] For instance, Bound and Stinebrickner (2010), Gallipoli and Turner (2011), Kitao (2014) , Low and Pistaferri (2015), Michaud and Wiczer (2018), Autor et al. (2019), Kellogue (2021).

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

[13] For instance, Flinn and Mullins (2015), Blundell et. al. (2016), Bobba and Flabbi (2018).

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 decision taken before entry into the labour market. If the labour market distortions created by SA or DI are large enough to significantly impact the returns to schooling for this group, they may also affect their chosen level of education. My research also relates to the literature studying how individuals make their education decisions given future uncertainty.[14] My contribution is to evaluate the role of SI policy in partially insuring against uncertainty and affecting expected values to education levels.

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 the these data. Section 5 details the empirical model. Section 6 discusses the strategy to estimate the model's parameters. Section 7 reviews the empirical results and Section 8 performs counterfactual experiments using the model. Finally, Section 9 concludes.

2 Disability Policy Environment in Canada

The Canadian SI environment is made up of a set of programs at both the provincial and federal levels, many of which are administered separately. 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 2017). 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 each other, rather than being jointly administered or unified in eligibility, as is the case in other countries. While this feature has been criticized for leaving gaps in social assistance for individuals with disabilities, it is advantageous from an econometrician's perspective because the effect of the parameters from a given program on behaviour is less confounded by parameters from other policies integrated with this program.

This paper focuses on the two main programs providing long-term income assistance and replacement for individuals affected by disability.[15] These are the Canadian Pension Plan Disability (CPP-D), the federal DI program, and provincial SA programs, which offer means-tested welfare payments.[16] The Canadian Pension

[14] For instance, Carnerio, Hansen, Heckman (2003), Cunha, Heckman, and Navarro (2005).

[15] 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 individual's in my data.

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

Plan (CPP) is the federal retirement pension program, and also 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 individuals that are 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 disability is prolonged if it is expected to be indefinite or likely to result in death.[17] The severity of the disability concerns the applicant's ability to engage in "substantially gainful activity" in the labour market. Substantially gainful is subjectively determined based on an applicant's perceived productivity in the labour market given the barriers imposed by their disability. 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. Adjudicators incorporate an individual's personal characteristics when determining an individual's capacity for substantial gainful activity. Most notably, personal characteristics include an individual's age, education, and work experience (Government of Canada 2018b).

The second main DI eligibility requirement for applicants is to have has made contributions to the CPP in four of the previous six years.[18] 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 generosity of CPP-D is calculated with an earnings index that summarizes the number of months in the contributory period and the size of the individual's monthly CPP contributions. In the calculation of the earnings index, applicants can drop certain months from their contributory period that may reduce their final amount of CPP benefits.[19] CPP-D payments are the sum of two components. The first component is

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

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

[19] 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.

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 applicants 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 each year based on a measure of average wages. The second component is a deterministic flat-rate benefit, which is indexed by the CPI each year.[20]

2.2 Provincial Social Assistance in Canada

The main source of welfare in the social safety net for Canadians comes from 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 intended to be a last resort, it is available for individuals who have exhausted all other means of assistance. This means 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 covered by these programs.

SA is separately administered in each province. As such, the SA programs vary in eligibility criteria and the generosity of their transfer across provinces. The SA programs all share a similar structure (Government of Canada 2017).[21] 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 their eligibility. An applicant is deemed eligible if their assessed needs exceed the sum of their income and assets, up to an upper threshold. First, an applicant's "needs" may include variables like living expenses, family size and composition, and disability.[22] Assessed income combines all earnings from market activities, such as paid employment or self-employment, with transfers from other government

[20] In 2018, the average CPP-D benefit received was just under \$1000 per month, half of which was the deterministic flat rate component (Government of Canada 2018a).

[21] 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), Beland (2015), Torjman and Makhoul (2016).

[22] 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, 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 Government of Canada (2017) or Hillel (2020).

programs, such as DI. [23] 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, clothes. The cost of living varies with an individual's demographic characteristics, notably their disability status. Thus, the basic assistance amount varies with a recipient's demographic characteristics. 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. I refer to these added resources as SA-Disability (SA-D).

3 Data: The Longitudinal and International Study of Adults

To study the relationship between SI and education, I use LISA, which 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 notably are T1 family files (T1FF), that 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 for my sample between 1989 and 2017. Finally, LISA includes a relatively large set of early-onset individuals with linked T1FF tax data and is, therefore, especially well-suited to this research agenda.

I use the survey waves of LISA to obtain the majority of demographic information used in this study. Each survey wave measures details 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. Education level is derived by the self-reported highest certificate of completion. I define a low education level as completing at most a high school certificate. The high education level includes those with any completed post-secondary education.

[23]The inclusion of unearned income is due to SA being designed as assistance of last resort.

Measuring Disability

The 2014, 2016, and 2018 waves of LISA include measures of LADLs and other characteristics of health, that are used to understand the history of disability status.[24] LISA includes a set of LADLs, which 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). There are five main distinct areas of activity limitation: Seeing, Hearing, Physical, Cognitive, and Mental Health. Physical includes any limitation to mobility, flexibility, dexterity, as well as limitations caused by pain. Cognitive includes learning, developmental and memory limitations. Mental Health includes anxiety, PTSD, depression, and other mental health conditions.

The activity limitations are self-reported, and the age of disability onset is retrospective. For each type of activity limitation, respondents were asked about the magnitude of difficulty and the frequency of limitation.[25] Individuals are flagged as disabled if they report to have any type of LADL and respond "sometimes", "often", or "always", to the frequency of their LADL.

Using self-reported functional limitations to measure disability is not without its share of criticism, as are all other methods of defining disability.[26] Opponents of using self-reported disability are often concerned with the endogeneity of reporting to ones labour market situation and over-reporting. However, 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). Using specific health questions, such as the activity limitation screening questions in LISA, rather than directly asking about disability status can further reduce the concerns associated with self-reports being endogenous to labour outcomes (Bound and Burkhauser 1999). The combination of this evidence is sufficient to validate the use of self-reported disability, as defined by specific questions on functional limitations.

T1 Family Files

The income histories come from the T1FF records. These data include demographic characteristics relevant to tax filings, such as age, marital status, and province of residence. The tax records span from 1982 to 2017 and are linked to each respondent in the main survey waves of LISA. I focus on 1989-2017

[24] The 2012 wave comprises only a small set of questions about the disability. Notably, the 2012 wave excludes the variable that determines the age of disability onset.

[25] Some cognitive conditions, such as developmental disabilities or learning conditions, were derived based on diagnosis from medical professionals instead of the level of difficulty. Refer to Appendix 9.1 for details.

[26] For instance, using DI beneficiaries to flag people with disabilities has been found to under-represent the population of individuals who are limited enough in the labour market to be classified as "disabled" (Bound 1989).

calendar years due to limitations on the income measures in the years prior.[27] of the years in the T1FF for longitudinal yearly measures of an individual's personal income. For this analysis, I focus on measures of employment income, and government transfers from CPP-D and SA payments. 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.[28]

Sample Selection

My sample of interest consists of males with and without an early-onset disability. I restrict the sample individuals between the ages 18 to 65 during the calendar years 1989 to 2017. LISA excludes institutional residents, whose disability eliminates any hope for participation in the labour market.[29] I exclude individuals living in the Canadian Territories.[30]

3.1 Social Assistance and Taxes

Data on the parameters of SA program are found in The Maytree Foundation's annual report series, "Welfare in Canada" (Maytree 2018).[31] This resource calculates the maximum amount from SA that a household may receive in their respective province and calendar year.[32] The calculated maximum amount of welfare also includes provincial and federal tax credits that may be available to individuals. The maximum amount of SA is 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.

The parameters of the Canadian income tax system using data from the Canadian Tax and Transfer Simulator, created and distributed by Kevin Milligan.[33] This rich resource reports income tax brackets and respective marginal tax rates for the federal and provincial level, 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.

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

[28] Meyer, Mok and Sullivan (2009) find survey reports on public transfers often suffer from respondents under-reporting, which can lead to overestimation of total income declines following the onset of disability.

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

[30] In the current state of this paper, I not vetted sample sizes to prevent restrictions to future vetting.

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

[32] The TIFF's include an individuals 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 are savings and assets.

[33] Details on this resources can be found on Kevin's website, https://sites.google.com/view/kevin-milligan/home.

4 Motivating Statistics

I next review a set of descriptive statistics from LISA to motivate the relationship between education and SI. Also, these descriptives are informative measures of the costs of an early onset disability. 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. The expected labour market gains to earnings and employment are a key component of the return to education. Disability interacts with human capital to determine productivity, and therefore, is linked to earnings and employment.

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

	Data
Early-Onset	0.4600
	(0.0370)
Not Early Disabled	0.6400
	(0.0120)

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.

The returns to education are seen when comparing outcomes across education levels. The difference in the returns to education for individuals with an early-onset disability is reflected in differences across education and early disability status. The first row of Table 2 contrasts the difference in earnings by early disability status and education level. The difference in annual earnings by education between early-onset and not early disabled individuals is reflective of the financial gain from accumulating human capital, keeping in mind that selection into higher education depends on unobserved factors, such as ability and motivation.[34] The gradient in annual earnings with respect to education for early-onset individuals is three-fourths that of their not early disabled counterparts in the data. This finding is consistent with a disability impeding the accumulation of human capital during school.

The difference in employment rates by education level and early disability status, shown in row 2, reflects differences in offered and reservation wages. Employment rates may also reflect labour market frictions, such as the rate of job offers or job loss, that affect the likelihood of employment. The annual rate of employment over the life-cycle is lower for early-onset, regardless of education level. These statistics suggest

[34]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.

Table 2: Employment, Earnings, and Labour Market Uncertainty 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($)	32300	50900	26000	40400
	(100)	(100)	(500)	(600)
Employment Rate	0.7400	0.8460	0.5080	0.7530
	(0.003)	(0.002)	(0.011)	(0.009)
First 3 years in Labour Market				
Annual Earnings ($)	15100	20700	12900	18200
	(300)	(230)	(600)	(800)
Employment Rate	0.810	0.862	0.579	0.815
	(0.009)	(0.006)	(0.028)	(0.022)
Labour Market Frictions (All Years in Labour Market)				
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.

that early-onset individuals may expect a lower return to employment and earnings from post-secondary, lowering the incentive to invest in their education.

Rows 3 and 4 of Table 2 compare average annual earnings and employment rates in the first three years of working life. Initial earnings are discounted the least in an individual's expectation of the return to education at the time of their schooling decision. At the start of working life, early-onset individuals experience a larger earnings gradient with respect to education. As the gradient for lifetime earnings with respect to education level is smaller, this finding suggests early-onset individuals experience relatively lower earnings growth with a post-secondary degree.

Observed employment rates and earnings reflect both the labour supply of individuals and the labour demand of employers.[35] The last panel in Table 2 compares the rate of job offers for individuals that 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 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 be a result of a higher lifetime dependence on government transfers for these individuals.

The culmination of this evidence motivates a role for SI policy in education decisions for early-onset individuals. 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 and in need of SI. Furthermore, this group enters working life with a disability, granting them more generous SI transfers. DI and SA are are most relevant for early-onset individuals with low education, as this group is most exposed to low and volatile earnings, higher employment risk, and health risks. Table 3 shows the likelihood of benefit receipt and average benefit of DI and SA for early-onset and not early disabled individuals.

In rows 1 and 2 of Table 3, there exists a stark difference 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 to 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 3 show the probability of becoming a recipient to CPP-D is low, culminating to

[35] Information on employers is limited in the T1ff files. However, the survey waves of LISA identifies 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.

Table 3: 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 $\geq$ 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 $\geq$ 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.

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 (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 at risk of requiring some form of government assistance in their life, especially for 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.

5 Empirical Life-Cycle Model

The descriptive evidence shows a correlation between SA, DI, and education. To analyze the factors driving this relationship, I consider a life-cycle model of education investments and labour market decisions 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, $d_0 \in \{0, 1\}$, and latent heterogeneity, $a_i \sim N(\bar{a}^{d_0}, \sigma^2_{a^{d_0}})$. Latent heterogeneity can be interpreted as an individual's "ability" and interacts with education to determine human capital at labour market entry. Initial disability status distinguishes the group of individuals with an early-onset disability, where $d_0 = 0$ corresponds to not early disabled and $d_0 = 1$ corresponds to early-onset.

Education level is chosen to maximize an individual's expected discounted lifetime utility. The expectation depends on a labour market environment and 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. 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.

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 previous decisions. Individuals chose to work, 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 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. Disability risk is assumed exogenous to individuals labour market decisions.[36] Disability transition probabilities vary with age and initial disability status. The risk of disability onset is increasing with age and the chance of recovery decreases with age. 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 is

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

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 may be exogenously displaced out of employment with probability, $\delta^{d_0,s}$. Employed individuals may choose to quit their job endogenously. These frictions are allowed to vary by early-onset disability status to account for differences in institutional features, employer beliefs, and other barriers to work for this group.[37]

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, $\epsilon_{it}^{d_0,s}$,

[36] Modelling disability risk this way is standard in the related literature (Low and Pistaferri 2015, Autor et al., 2019, Michaud et al., 2017, Kellogue 2021).

[37] Some resources that discuss how these features may lead to different labour market frictions for individuals with disabilities are Acemoglu and Angrist (2001), Dixon et al., (2003), Kitao (2014), Morris et al (2017).

and unobserved fixed heterogeneity, $v^{d_0}(s_i, a_i)$. Potential earnings in an arbitrary period, t, are

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

$$\text{where } \epsilon_{it}^{d_0,s} = \epsilon_{it-1}^{d_0,s} + \xi_{it}^{d_0,s},$$

$$\xi_{it}^{d_0,s} \sim N(0, \sigma^2_{\xi^{d_0,s}}) \text{ for } t > 0, \tag{3}$$

$$\text{and } \xi_{i0} \sim N(\bar{\xi_0}, \sigma^2_{\xi_0}).$$

The parameters governing potential earnings depend on initial disability status and education level. This specificity of parameters allows for differences in the life-cycle evolution of earnings. 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 allows early disability to affect 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.[38] The direct effect of a disability on productive human capital is captured by ϕ. This encompasses a disability induced loss of work-relevant skills, shifting earnings growth downward.

Permanent productivity shocks, $\epsilon_{it}^{d_0,s}$, follow a random walk with iid 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 would impact 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:

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

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.[39] Initial earnings also depend on an initial human capital shock,

[38]Cutler, Landrum and Stewart (2006) study heterogeneity across education levels in ones ability to cope with a disability affecting the evolution of their life-cycle earnings.

[39]This specification is similar to Flinn and Mullins (2015) and has the feature that earnings production is supermodular in

ϵ_{i0}, representing productive human capital that is unrelated to education, such as interpersonal skills.

The Earnings Index

The earnings index, which is a summary measure of an individual's earning history in T^L, is used to determine the generosity of DI and retirement transfers.[40] The earnings index, e_t, is assumed to update each period given the previous periods earnings index, e_{it-1}, the individual's labour earnings in the current period, W_{it}, and age, t, according to the function *f*. The function f for $t \in T^L$ is:

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

and $e_{it} = 0$ for $t \in T^S$. The variables $\underline{W}$ 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 are approximately the real value of the upper and lower bound in reality. 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.[41] Retirement income comes from pension benefits and old age security. Retirement benefits equal $0.25 * e_{it}$, which approximates is the formula for 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.[42] OASP helps supplement income for those with no CPP income.

Individuals may be 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 at age 60. The penalty lasts for the duration of their retirement, and individuals do not receive \$5,500 from OASP until age 65.[43]

ability. $h_0^{d_0}$ is normalized to 1, so individuals mean ability combines their ability endowment with human capital production before age 18.

[40] 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.

[41] Education and early-onset disability may affect determinants of early retirement. For instance, through affecting health at older ages or by influencing the accumulation of wealth.

[42] This value is consistent with average OASP income observed in LISA.

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

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 some simplifying assumptions ot the DI program for computation tractability. 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 modelled 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.

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 zero otherwise. I assume DI is only available for those with $d_{it} = 1$.[44] Hence, conditional on having applied to DI in the previous period, $m_{it-1} = 1$, the probability of acceptance is

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

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.[45] The yearly flat rate component of DI of \$4,365 is set by policymakers and known to agents in the model.[46] Hence, DI generosity is given by

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

I also fix the maximum amount to \$40,000 and the yearly basic exemption is \$3,500, which is representative of the true program.[47]

Individual's face a utility cost of applying to DI, $C_{App}^{d_0,s}$. The application process can be lengthy and

[44]As disability is measured based on limitations to daily activities, this assumption may miss some individuals with a 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 minuscule and only occurs at the very end of the life-cycle for this group.

[45]I model DI generosity in a similar manner as Gallipoli and Turner (2011) and Milligan and Schirle (2019).

[46]The real value of this amount has fluctuated from \$3,900 - \$4,500 and I fix it to in my model, which is the average over the years covered by my sample. This is in real dollars with a base year of 2002.

[47]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.

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 individuals 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 preference 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 in order to become a beneficiary again.

Social Assistance

SA benefits are means-tested anti-poverty programs. The maximum amount of benefits from SA programs differ 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. 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}. Define $\mathbb{1}^{SA-D} = 1$ if $d = 1$ and approved for SA-D. Then,

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

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 years and provinces observed in my data.

Preferences

I assume individuals have a non-separable CRRA utility specification.[48] 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 is consistent with disability and work needing higher consumption to have the same utility as non-disability or non working.[49] These parameters capture the utility loss induced by work and disability, respectively. The coefficient of risk aversion is greater than 1 so that individuals are risk averse, which is an important assumption when studying the effect of social insurance on behaviour.[50]

Individual's Problem in the Labour Market

These individual features and market environment combine to define an individual's decision problem 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:

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

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

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

Agents decide to work or apply for DI in order to maximize their discounted lifetime utility, equation (10), subject to budget constraint, (11), and the evolution of their earnings index, (12).[51] Utility from future periods is discounted by β. The expectation operator, $\mathbb{E}_t$, is conditional on the set Ω_t, which includes individual

[48] This specification for preferences has been adopted in various related studies, notably, Low and Pistaferri (2015), Michaud and Wiczer (2018) and Autor et al. (2019).

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

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

[51] I do not model savings decisions due to data limitations. A justification for excluding this this modeling choice is that my population of interest, individuals with disabilities whose decisions are influenced by DI, do not tend to earn enough to have significant savings (Bound and Stinebrickner 2010). However, there are parameters that are estimated off the entire sample, many of whom would have savings. Savings act as self-insurance, so the assumption of full take-up of SA when not working will replace some of the role for savings.

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 are 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, $\tau()$, which represents the Canadian combined provincial and federal tax system.[52] 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 in terms of utility. Education investments may be more costly in the presence of disability. Individuals face an 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, $\epsilon_\psi \sim N(0, \sigma_{\epsilon^\psi})$.[53] 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

[52] 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.

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

$S_t = \{d_t, \epsilon_t, e_{t-1}, elidg_t\}$.[54] Given the solution to the labour market behaviour, 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, $S_t = S_{t+1} = \bar{S} = \{\bar{d}, \bar{\epsilon}, \bar{e}, \bar{elidg}\}$. 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.[55] I assume individuals expect retirement to last until age 75, after which they die with certainty. The value in 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 update to ϵ_{t+1} and d_{t+1}. 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 to leave work. The value function for employed individuals is

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

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

[55]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 they are rejected, they do not receive a job offer, and they remain out of work 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, or to remain out of work. If the individual does not receive a job offer, they remain out of work next period. The value function for an unemployed individual at age t is

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

$$\begin{aligned} \text{s.t. } & c_t = SA(0, d_t), \\ & e_t = f(e_{t-1}, 0, t). \end{aligned}$$

DI Beneficiary

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

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

$$\begin{aligned} \text{s.t. } & c_t = \tau\big(0,\ DI_t\big) + SA_t(\tau\big(0,\ DI_t,\big), d_t) \\ & e_t = e_{t-1}. \end{aligned}$$

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 chooses to go to post-secondary if

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

The inequality in equation (18) captures how early disability may influence educational investments through 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 individuals on the margin.

6 Model Estimation and Identification

Identification of the model's parameters consists of two stages. In the first stage, a set of parameters are calibrated to values from the related literature or are estimated external to the life-cycle model. In the second stage, the remaining parameters are estimated via indirect inference given the parameters 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.

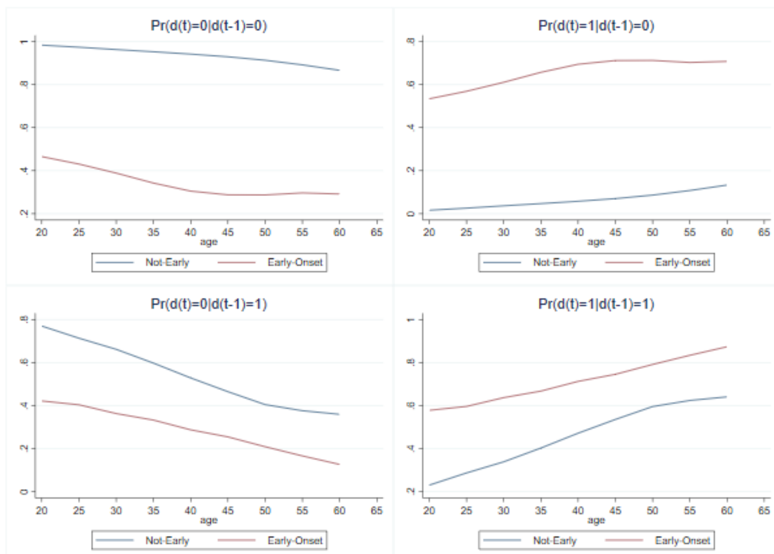
First, I set the coefficient of risk aversion to the value used in Low and Pistaferri (2015) and in Blundell et. al (2016), $\kappa = 1.5$. This ensures individuals are risk averse in the model. The discount factor is calibrated to the value in Low and Pistaferri (2015) $\beta = 0.9756$.[56] I take the 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 (Office of the Chief Actuary 2011).

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

Disability Risk

Under the assumption of disability risk being exogenous to the choices of agents in the model, I estimate the disability transitions using observed transitions in the survey waves of LISA.[57] 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.[58] 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



The top right panel of figure 1 shows the probability of inuring a disability shock increases with age and is much larger greater for early-onset individuals. This is consistent with a high rates of disability re-occurrence of disability in adulthood for those with a disabling condition early in life. 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.

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

[57] Disability status is unobserved in the tax records. I am implicitly assuming the transition probabilities in the over the calendar years covered by LISA, 202-2018, are representative of the calendar years covered by the tax records, 1989-2018.

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

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.[59] 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 set of estimated parameters, $\hat{\Theta}$, are defined by

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

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 weight each mean squared difference using the variance of M_{kN}^{d}, $Var(M_{kN}^{d})$.[60] An exception are moments capturing the distribution of post-secondary, which I use an order of magnitude smaller than the variance of the data moment.[61] I have

[59] 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.

[60] Using $Var(M_{kN}^{d})$ to weight the moments is conventional in the related literature. The asymptotically optimal weight matrix has potentially poor small sample properties (Altonji and Segal 1996).

[61] This decision helps to match the schooling distribution in the data substantially. However, this decision comes at the expense of the efficiency of the parameter estimates.

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). The distribution of early-onset disability is representative of the distribution in LISA.

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

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

This model is estimated separately by d_0 and s. Notice, this model is similar to the sample analogue of the first difference of the earnings process (2), 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.[62]

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

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

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 pin down these 7 parameters. To supplement these moments, I also compute the mean and variance of annual earnings in the first 3 periods in the labour market, conditional on s and d_0. The first three periods are affected by productivity shocks the least andso 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

[62] 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.

absorbing. I estimate the residual,

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

separately by s and d_0 to pin down $\sigma^2_{\xi^{d_0,s}}$.

To further help recover the true earnings profile over the life-cycle, I use the OLS estimates from the following conditional regression:

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

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. The variance of measurement error is identified off the variance and auto covariance of a residual from the following model.

$$LnW_{it} = \beta_0 + \beta_1 PE + \beta_2 PE^2/100 + \beta_4 LnW_{it-1} + u_{it} \quad (1)$$

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:

$$\mathbb{1}(DI_{it}) = \beta_0^{s,d_0} + \beta_1^{s,d_0} t + \beta_2^{s,d_0} t^2 + \beta_3^{s,d_0} t^3, \text{ and} \tag{2}$$

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

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.

Utility Cost of Working.

To estimate the utility cost of working, I match employment rates over the life-cycle conditional on (s, d^*, d_0) and age greater than or less than 45. These moments are informative of the utility cost of working, η. In addition, I match the flows into and out of employment by the same conditioning variables.

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. $\hat{v^{d_0}}(a_i, s_i)$ 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 4 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 4, 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 greater range of human capital at the end of high school for early-onset individuals.[63]

Table 4: Estimates of Parameters for Individual Heterogeneity and the Earnings Process.

		Not Early Disabled		Early-Onset	
		Low Education	Post-Secondary	Low Education	Post-Secondary
Conditional Parameters					
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).

Row three shows that attending post-secondary scales initial ability by 3.1% for early-onset individuals

[63] 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.

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.

Table 5: Estimates of Remaining Parameters

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).

The remaining structural parameters are reported in Table 5. First, the utility cost of working, η, equates to approximately 8% of annual consumption. The likelihood of acceptance to DI, shown in the third and fourth rows, is modestly lower for applicants with post-secondary.[64] 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 (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 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

Next, I compare the fit of the estimated model relative to 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 can help validate the parameters and help with their interpretation. First, Table 6 compares 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 school is the financial gain that results from investing in productivity. Table 7 reports moments of employment earnings by early disability status and education level. The first row presents average yearly earnings over the life-cycle. The simulated model recovers an average financial return to education of \$6,410 for early-onset individuals. This difference is attenuated relative to the data, which observes an average financial return to education of \$14,400. This attenuation is mainly caused by the average annual earnings of early-onset individuals with low education being driven up

[64] 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 parameter specification.

Table 6: 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.

Table 7: Simulated Earnings Moments Relative to Moment in the Actual Data

		Not Early Disabled		Early-Onset	
All years in the Labour Market		Data	Sim	Data	Sim
Average	Low Education	32300	34728	26000	32095
		(100)		(500)	
	Post-Secondary	50900	51529	40400	38505
		(100)		(600)	
Q10	Low Education	9400	11029	5200	9252
		(132)		(374)	
	Post-Secondary	13700	15244	9200	10827
		(151)		(570)	
First 3 Years in Labour Market					
Average	Low Education	15100	15308	12900	10064
		(300)		(600)	
	Post-Secondary	20700	18864	18200	16201
		(230)		(800)	

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

by a small number of high earning outliers.[65] The bottom panel of Table 7 shows average yearly earnings in the first three years of working life. Expected earnings at labour market entry are the least discounted when individuals make their education choices. The model does a good job of matching the level of initial earnings for all groups. The model recovers an initial earnings premium to post-secondary of \$6,137 for early-onset individuals, compared to \$5,300 in the actual data.

[65] This group faces the largest estimated idiosyncratic shock to their productivity, ξ. The average simulated yearly earnings equals \$27,302 when excluding the top 5% of simulated earnings.

The middle row in Table 7 reports the cutoff of the bottom decile of the distribution of annual earnings over the entire life-cycle. Individuals in the bottom decile of the earnings distribution are most likely to require assistance from SA or DI. Again, the estimated model recovers a positive earnings premium to education that is attenuated compared to the premium observed in the actual data.

Lastly, Table 8 presents rates of employment and DI over the life-cycle. Focusing first on the top

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

		Not Early Disabled		Early-Onset	
Employment Rate		Data	Sim	Data	Sim
Low Education	$age < 45$	0.8743	0.8456	0.5213	0.6269
		(0.0030)		(0.0140)	
	$age \geq 45$	0.7974	0.8283	0.4799	0.5954
		(0.0040)		(0.0200)	
Post-Secondary	$age < 45$	0.9076	0.9292	0.8152	0.7702
		(0.0020)		(0.0090)	
	$age \geq 45$	0.8504	0.9540	0.6107	0.6843
		(0.0030)		(0.0180)	
Rate of DI					
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.

panel, I partition employment rates by early disability status, education level, and age. The structural model recovers a positive employment gradient with respect to education for both early-onset and not early disabled individuals. Furthermore, the model predicts a decline in employment over the last 20 years of working life for all, except for individuals not initially disabled with post-secondary. The level of employment is simulated higher in the model relative to the data for both education levels. The consequence of this is that the rate of SA is under-simulated, which must be kept in mind in the policy experiments.[66] Finally, in the bottom panel of Table 8, I report the rate of DI 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.

[66] In reality, the value of not working may be higher from access to other resources, such as personal savings, which may explain lower employment rates in the 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. Lastly, I consider the life-cycle effects of providing early-onset individuals with a grant that subsidizes their consumption during post-secondary.

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 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.[67]

This decomposition method has three desirable features. First, the decomposition is exact, meaning each factor's contribution sums to equal the baseline gap.[68] Second, this decomposition is symmetric, which occurs when the estimation of the each factor's contribution is path independent to the order estimating other factor's contributions.[69] Lastly, this method accommodates hierarchical structures within the contributing factors. I conduct a primary decomposition, where the contribution of the primary factors can be

[67]For a step-by-step breakdown of this procedure, see Shorrocks (2013)

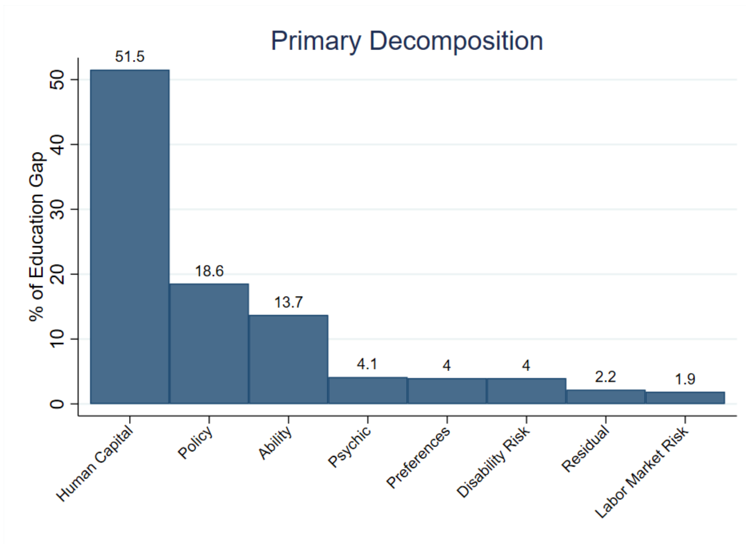
[68]An example of a non-exact decomposition is to set the contribution of each factor according to turning that factor off with all others turned on. This kind of decomposition for this model can be found in Appendix.

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

decomposed into the contribution of a set of secondary factors. I consider seven primary factors, three of which have a secondary structure.

Figure 2 plots the results from the primary decomposition. From left to right, the seven primary

Figure 2: Decomposing the Education Gap: Primary Decomposition



factors are human capital, policy, individual ability, the psychic cost to schooling, the effect of disability on preferences, difference 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 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 3: Decomposing the Education Gap: Secondary Decomposition

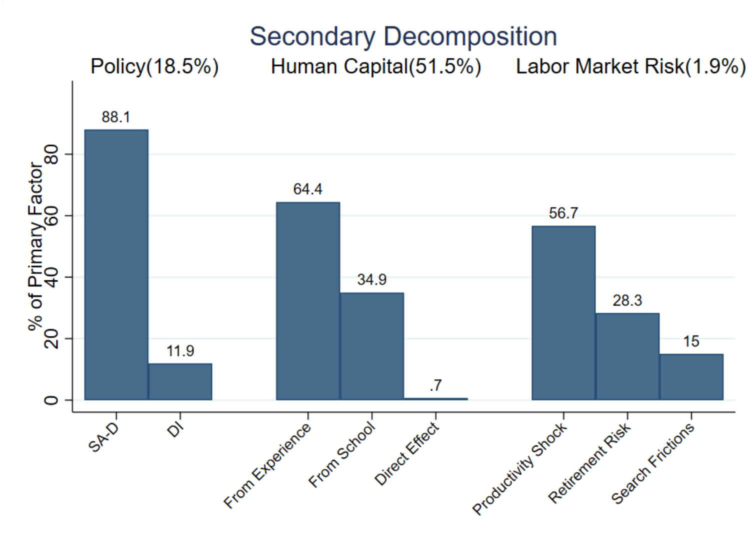


Figure 3 reports the results from the secondary decomposition. The primary factors for policy, human capital, and labour market risk have a secondary structure. First, the education-specific returns to experience account for two-thirds of the human capital factor. Early-onset disability impedes the accumulation 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.[70]

Labour market risks don't have a large 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.

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

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 policies that achieve said objective. I fix the policy objective of increasing post-secondary completion of early-onset individuals by one percentage point,(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 an insurance-incentive trade-off.

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 so they have better job opportunities when the subsidy expires.

Figure 4: policy on change in government revenue and liabilities.

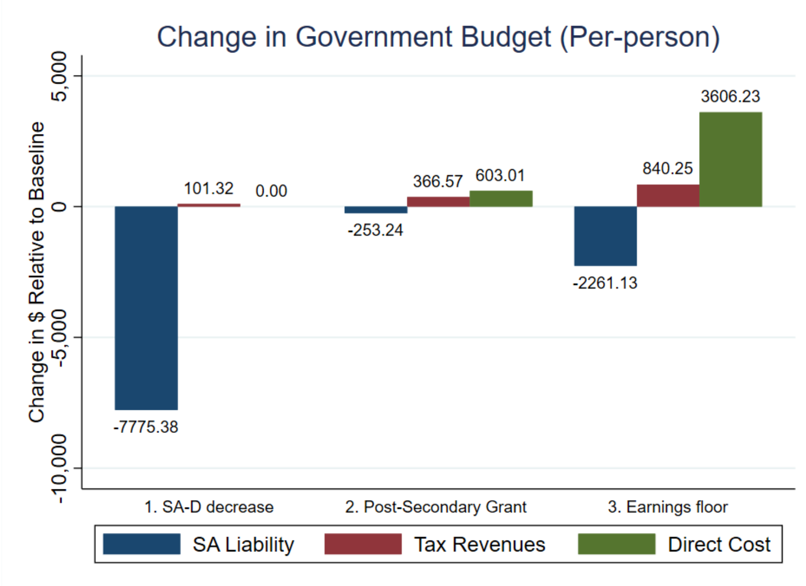
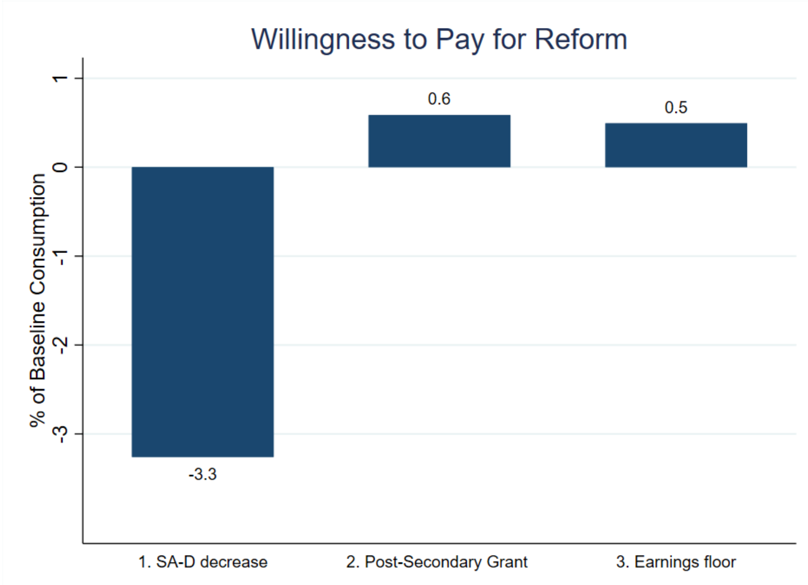


Figure 4 summarizes the change in government revenue and liability relative to the baseline for each policy reform. The bars are the difference in average lifetime present value dollars per person. The blue bar shows the change in SA liability, the red bar indicates the change in tax revenues, and the 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 5 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 implement the policy change in terms of a proportional reduction to current and future consumption.[71] 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.

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

Figure 5: Policy reforms on willingness to pay



Unsurprisingly, taking resources away from people at the left tail of the income distribution greatly reduces their welfare.[72] 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.

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, notable SI policy. 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

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

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 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 towards enhancing individuals' economic independence.

10 References

Acemoglu, Daron, and Joshua D. Angrist. 2001. Consequences of Employment Protection? The Case of the Americans with Disabilities Act. Journal of Political Economy 109(5): pp.915-57.

Almond, D. and Currie, J., 2010. Human Capital Development before Age Five. NBER Working Paper No. 15827. National Bureau of Economic Research.

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

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

Attanasio, Orazio P., James Banks, Meghir Costas, and Guglielmo Weber. 1999. Humps and Bumps in Lifetime Consumption. Journal of Business and Economic Statistics.

Attanasio, Orazio P., and Guglielmo Weber. 1995. Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey. Journal of Political Economy 103(6): pp.1121-57.

David, H. and Duggan, M.G., 2006. The growth in the social security disability rolls: a fiscal crisis unfolding. Journal of Economic perspectives, 20(3), pp.71-96.

Autor, David, Andreas Kostøl, Magne Mogstad, and Bradley Setzler. 2019. Disability Benefits, Consumption Insurance, and Household Labor Supply. The American Economic Review 109(7): pp.2613-2654.

Banks, James, Richard Blundell, and Agar Brugiavini. 2001. Risk Pooling, Precautionary Saving and Consumption Growth. Review of Economic Studies 68(4): pp.757-79

Béland, D., & Daigneault, P.-M. 2015. The state of provincial social assistance in Canada (1442609710). Retrieved from https://www.schoolofpublicpolicy.sk.ca/documents/research/policy-briefs/JSGS-policybriefs-Social-Assistance_WEB.pdf

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

Bobba, M., Flabbi, L. and Levy, S., 2018. Labor Market Search. Informality and Schooling Investments, Toulouse School of Economics Working Paper, 863.

Bound, John. 1989. The Health and Earnings of Rejected Disability Insurance Applicants. The American Economic Review 79(3): pp.482-503.

Bound, John, and Richard V. Burkhauser. 1999. Economic analysis of transfer programs targeted on people with disabilities. Handbook of Labor Economics 3: pp.3417-3528.

Bound, John, Todd Stinebrickner, and Timothy Waidmann. 2010. Health, Economic Resources and the Work Decisions of Older Men. Journal of Econometrics 156(1): pp.106-129.

Bruins, Marianne, James A. Duffy, Michael P. Keane, and Anthony A. Smith Jr. 2018. Generalized indirect inference for discrete choice models. Journal of Econometrics 205(1): pp.177-203.

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), pp.251-269.

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), pp.541-555.

Cagetti, M., 2003. Wealth accumulation over the life cycle and precautionary savings. Journal of Business & Economic Statistics, 21(3), pp.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. and Paxson, C., 2010. Causes and consequences of early-life health. Demography, 47(1), pp.S65-S85.

Cossette, Lucie, and Édith Duclos. A profile of disability in Canada, 2001. Statistics Canada, Housing Family and Social Statistics Division, 2002.

Cunha, F., Heckman, J. and Navarro, S., 2005. Separating uncertainty from heterogeneity in life cycle earnings. Oxford economic papers, 57(2), pp.191-261.

Currie, Janet. 2009. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. Journal of Economic Literature 47(1): pp.87-122.

Cutler, D.M., Landrum, M.B., Stewart, K.A. and Hurd, M.D., 2009. 6. How Do the Better Educated Do It? Socioeconomic Status and the Ability to Cope with Underlying Impairment. University of Chicago Press, pp.203-254.

Dixon, Karen A., Douglas Kruse, and Carl E. Van Horn. 2003. Restricted Access: A Survey of Employers about People with Disabilities and Lowering Barriers to Work. New Brunswick, NJ: Rutgers University, John J. Heldrich Center for Workforce Development.

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

Gallipoli, Giovanni, and Laura Turner. 2009. Household responses to individual shocks: Disability and labor supply. No. 831-2016-55359.

Gourieroux, C., Monfort, A. and Renault, E., 1993. Indirect inference. Journal of applied econometrics, 8(S1), pp.S85-S118.

Gourinchas, P.O. and Parker, J.A., 2002. Consumption over the life cycle. Econometrica, 70(1), pp.47-89.

Government of Cannada. 2017. Social Assistance Statistical Report: 2009-2013. Employment and Social Development Canada. Retrieved from: https://www.canada.ca/en/employment-social-development/services/social-assistance/reports/statistical-2009-2013.html

Government of Canada. 2018a. Annual Report of the Canada Pension Plan 2017–18. Her Majesty the Queen in Right of Canada.

Government of Canada. 2018b. Canada Pension Plan Adjudication Framework. Retrieved from Government of Canada Website: https://www.canada.ca/en/employment-social-development/programs/disability/benefits/framework.html#h2.3.

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

Haveman, Robert, and Barbara Wolfe. 2000. The economics of disability and disability policy. Handbook of Health Economics 1: pp.995-1051.

Hillel, I., 2020. Holes in the social safety net: Poverty, inequality and social assistance in Canada (No. 2020-06). Centre for the Study of Living Standards.

Keane, M.P., Todd, P.E. and Wolpin, K.I., 2011. The structural estimation of behavioral models: Discrete choice dynamic programming methods and applications. In Handbook of labor economics (Vol. 4, pp. 331-461). Elsevier.

Kitao, Sagiri. 2014. A Life-Cycle Model of Unemployment and Disability Insurance. Journal of Monetary Economics 68(C):1-18.

Kellogg, M.D., 2021. Spousal Labor Supply and the Welfare Implications of Disability Insurance Reform.

Kneebone, R. and White, K., 2015. An overview of social assistance trends in Canada. Welfare reform in Canada: Provincial social assistance in comparative perspective, pp.53-92.

Low, Hamish, and Luigi Pistaferri. 2015. Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off. The American Economic Review 105(10): pp.2986-3029.

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

Maytree. 2018. Social Assistance Summaries 2018. Retrieved at:
https://maytree.com/wp-content/uploads/Social_Assistance_Summaries_All_Canada.pdf.

Meyer, Bruce and Mok, Wallace K. C. 2018. Disability, Earnings, Income and Consumption. Journal of Public Economics 171: pp.51-69.

Meyer, B.D., Mok, W.K. and Sullivan, J.X., 2009. The under-reporting of transfers in household surveys: its nature and consequences (No. w15181). National Bureau of Economic Research.

Michaud, A. and Wiczer, D., 2018. The disability option: Labor market dynamics with macroeconomic and health risks (No. 2018-7). CHCP Working Paper.

Milligan, Kevin. 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.

Millard. 2021. The Longitudinal Effects of Disability Types on Incomes and Employment. Working Paper. Retrieved from https://bobmillard.github.io

Mori, Hiroak., 2016. Essays on Human Capital Complementarities. Electronic Thesis and Dissertation Repository. 3898. https://ir.lib.uwo.ca/etd/3898.

Morris, Stuart, Gail Fawcett, Laurent Brisebois, and Jeffrey Hughes. 2018. A Demographic, Employment and Income Profile of Canadians with Disabilities Ages 15 Years and Over, 2017.

Office of the Auditor General of Canada. 2015. Report 6—Canada Pension Plan Disability Program. Retrieved from https://www.oag-bvg.gc.ca/internet/English/parl_oag_201602_06_e_41063.html#hd2c.

Office of the Chief Actuary. 2011. Canada Pension Plan Experience Study of Disability Beneficiaries. Actuarial Study No. 9. Retrieved from: https://www.osfi-bsif.gc.ca/Eng/Docs/cppas9.pdf

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

Shorrocks, Anthony F. 2013. Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. Journal of Economic Inequality 11, no. 1. pp.99-126.

Stern, S., 1989. Measuring the effect of disability on labor force participation. Journal of human Resources, pp.361-395.

Torjman, Sherri. 2002. The Canada Pension Plan Disability Benefit. Ottawa: Caledon Institute of Social Policy.

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

Tweddle and Aldridge 2018 Welfare in Canada, 2017. Maytree Foundation. Caledon Institute of Social Policy.

Appendix

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
$\gamma_{i,j}^{d_0,t}$	Disability risk for transitioning from $d_t = i, d_{t-1} = j$
Earnings Process	
ϕ	Direct effect of disability on earnings
$\mu_1^{d_0}$	Return to potential experience for d_0
$\mu_2^{d_0}$	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
$\pi_{ret}^{d_0,s}$	Retirement risk for s, d_0 at age t
C_{App}^{s,d_0}	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 $d_0 = 1$

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.

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 14

Table 10: 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

Auxiliary Moments

Tables 15 to 29 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 11: Post-Secondary Distribution

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

Table 12: 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 13: Employment Rates

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

Table 14: Initial Employment

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

Table 15: Flows Into and Out of Employment

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

Table 16: DI Rates, Flows, and Coverage

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

Table 17: 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 18: 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 19: Mean and Variance of Initial Earnings (conditional on working)

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

Table 20: 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
$\text{PE}^2/100$	-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
$\text{PE}^2/100$	-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
$\text{PE}^2/100$	-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
$\text{PE}^2/100$	-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 21: 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 22: 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 23: 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 24: 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 25: 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
$(1 - Fr(L_{it} = 1 \mid DI_{it+2} = 1))$	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
$(1 - Fr(L_{it} = 1 \mid DI_{it+2} = 1))$	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
$(1 - Fr(L_{it} = 1 \mid DI_{it+2} = 1))$	0.3280	0.6273	0.0360	0.2993

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 9.

Table 26: 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