Early-Onset Disability, Education Investments, and Social Insurance

By Robert Millard*

April 4, 2025

Abstract

This paper examines the impact of labour market social insurance policies on pre-market human capital investments, focusing on the post-secondary education gap between individuals with early-onset disabilities (occurring by age 19) and their non-disabled peers. Using a dynamic life-cycle model estimated with Canadian survey and tax data, I decompose the education gap, highlighting that policy-induced disincentives significantly reduce education rates for individuals with early-onset disabilities. Policy experiments show that reducing the generosity of social insurance lowers education disincentives and government costs but harms the welfare of individuals with early-onset disabilities. Alternatively, targeted consumption subsidies during post-secondary encourage investments and reduce government expenses without compromising welfare.

JEL Codes: I00, I21, I38, J14, H50 Keywords: Disability, Education, Social Insurance, Dynamic Disincentive, Welfare

^{*}Assistant Professor, Department of Economics, Stony Brook University, 100 Nicolls Rd, Stony Brook, New York, USA, 11794. Email: robert.millard@stonybrook.edu.

This paper uses confidential data that is administered by Statistics Canada and can not be made publicly available. Steps for data access can be found at https://www.statcan.gc.ca/en/microdata/data-centres/access. Supplementary files are provided to replicate the analysis provided access to these data. I would like to thank Audra Bowlus, Nirav Mehta, Christopher Robinson, and Todd Stinebrickner for their supervision of this project and helpful feedback and comments. Thank you to Rory McGee, Lance Lochner, Tim Conley, Sergio Ocampo Diaz, Al Slivinski, Elizabeth Caucutt, Salvador Navarro, Lisa Tarquinio, Baxter Robinson, David Rivers, Roy Allen, Neale Mahoney, Giovanni Gallipoli, Siha Lee, Timothy Moore, Steven Stern, David Wiczer, Hugo Benitez Silva, Marco Pariguana, and Tom Triveri for helpful feedback and discussions. I am also grateful for helpful comments from participants of the 2022 Canadian Economics Association Annual Meeting, 2022 Midwest econometrics group conference, 2023 Association for Education and Finance Policy conference, 2023 Society of Economics of the Household conference, 2023 International Association of Applied Economists, and seminar participants at Stony Brook University, Queens University, the University of Windsor, and the University of Western Ontario. This research is supported in part by funding from the Social Sciences and Humanities Research Council. Otherwise, the author declares that he has no relevant or material financial interests that relate to the research described in this paper. This article makes use of data administered by the Maytree foundation, Copyright 2023 Maytree. All rights reserved.

1 Introduction

Social insurance (SI) provides vital resources for individuals experiencing barriers to their gainful employment. However, this insurance comes with the well-known trade-off of disincentivising work and human capital investments in the labour market. In some cases, the distortionary incentives from SI may affect behaviour before labour market entry, particularly regarding education investments, but less is known about these dynamic disincentives. The behavioural incentives from SI are exacerbated for the population with disabilities, since disability can substantially lower the returns to work, and disability transfer payments are typically larger.¹ Understanding the insurance-incentive trade-offs for the disabled population is policyrelevant, as rising caseloads threaten the fiscal solvency of disability programs (Autor and Duggan, 2006; Autor, 2011; Liebman, 2015; Milligan and Schirle, 2019). Addressing the moral hazard related to dynamic disincentives is one way to alleviate financial liabilities by increasing the educational attainment of potential applicants and their subsequent returns to work.

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

This paper estimates a structural life-cycle model of post-secondary education investments and labour supply to quantify the dynamic disincentive associated with disability policy. The model accounts for a rich set of channels by which disability can discourage education to isolate the disincentives from policy. Education choices are dynamic in nature and made based on the expected returns to employability and earnings. First, education is an investment into one's productivity, and the expected return to schooling is increasing in one's human capital. A disability can interfere with productive skill development during childhood and adolescence, resulting in lower ability when post-secondary choices are made (Heckman, 2007; Currie and Almond, 2011). Additionally, a disability can interfere with skill accumulation during post-secondary and in the labour market afterward (Cutler et al., 2006; ?). Second, education is a costly investment, and the presence of a disability can exacerbate the financial and psychological costs of post-secondary.⁴ Finally, disability alters the labour market environment in ways that lower expected returns to work, discouraging schooling investments for forward-looking individuals. A disability can lower the expected likelihood of finding and

¹Disability 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)

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

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

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

maintaining employment.⁵ Moreover, SI will discourage education if it raises the expected outside option to work. This last channel is especially relevant for individuals with an early-onset disability, given their lower productivity and the greater generosity of SI programs available to them. Moreover, this group has the potential to rely on SI for longer durations, given the timing of their disability.

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

I estimate the model using the Longitudinal and International Study of Adults (LISA), a panel survey of Canadian households that contains rich information on health, education, and other demographic characteristics. LISA is linked to a 36-year panel of annual income tax records that contain disaggregated measures of personal incomes, taxes, and transfers. These administrative tax data reduce concerns of measurement error and under-reporting that are often associated with self-reported survey measures of income (Gallipoli and Turner, 2009; Meyer et al., 2009). The merged survey and administrative information facilitate the creation of a rich panel dataset with detailed health measures, demographics, and incomes. LISA provides data on life-cycle outcomes for a relatively large subsample of individuals with early-onset disabilities, which is uncommon in most panel studies.

The estimated model fits the data well in reproducing the gap in educational attainment, along with differences in life-cycle earnings and labour market trends by early disability status and education level. Early-onset individuals also have a lower financial return to education and work experience, representing the disruption of skill accumulation during post-secondary schooling and in the labour market. Moreover, the model reproduces life-cycle rates of DI receipt, and the estimated DI acceptance probabilities are consistent with the unconditional acceptance probabilities observed in the 2015 audit of the Canadian Federal DI program (Office of the Auditor General of Canada, 2015).

The estimated model is used to decompose the gap in post-secondary education between individuals with and without an early-onset disability and measure the role of dynamic disincentives. These exercises find disability policy to be a sizable contributor to the gap, which reduces by nearly 40% when eliminating disability programs in the model while holding all else fixed. SA-D is the main source of policy-related dynamic disincentives. SA-D benefits raise the relative value of low-earning states, causing individuals with

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

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

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

low potential earnings to substitute away from pursuing post-secondary education. DI holds little weight in the relative value of post-secondary education. This is because the expected value of DI is increasing in average labour market earnings, and individuals with early-onset disabilities have lower earnings on average. Moreover, the expected value of DI is heavily discounted at the time individuals make their schooling decisions, as people tend to apply for DI at older ages. Differences in the distribution of idiosyncratic costs, representing both pecuniary and non-pecuniary factors, are the primary contributors to the education gap. The impact of disability on skill accumulation in childhood and adolescence and skill accumulation during post-secondary are also large contributors to the gap.

Last, I use the estimated model to evaluate the effects of counterfactual policy reforms on the educational attainment and life-cycle behaviour of individuals with early-onset disabilities. Motivated by the decomposition analysis, I examine two classes of policy reforms aimed at increasing post-secondary attainment within this population. The first set of reforms involves proportional adjustments in the generosity of SA-D benefits, ranging from 80% to 120% of the baseline value. Reductions in SA-D generosity are found to increase post-secondary participation, employment, and earnings among early-onset individuals. These outcomes reflect a decrease in the relative value of not working, which raises the relative returns to education. As a result, government expenditures on early-onset disability beneficiaries decline, and tax revenues increase. However, these gains come at a substantial welfare cost as early-onset individuals are willing to forgo up to 5% of their lifetime consumption in the baseline to avoid a 20% reduction in SA-D generosity. These findings highlight the fundamental trade-off between insurance value of disability policy and education incentives.

The second class of reforms focuses on directly incentivising education through targeted consumption subsidies for early-onset individuals pursuing post-secondary education. By lowering the idiosyncratic costs faced by individuals with high potential ability, these subsidies promote educational attainment, which in turn leads to higher employment, increased earnings, and reduced dependence on disability programs, both SA and DI. Despite the direct fiscal cost of providing subsidies, the induced behavioural responses generate net government savings through higher tax revenues and lower liability to DI and SA. Though stylized, the two experiments give a transparent understanding of the insurance–incentive trade-off inherent in policy reform. Targeted consumption subsidies mitigate the welfare losses associated with reduced SA-D generosity while simultaneously proving to be fiscally efficient.

The contribution of this research is framed in three broad areas. First, I contribute to a sizable literature on the relationship between early-life health, education investments, and labour market outcomes.⁸ Health conditions at young ages can impede one's development in ways that persist for one's entire life. My contribution complements these studies by distinguishing and comparing the relative importance of mechanisms, such as human capital or the SI environment, underlying the education gap and analyzing how this relates to adult inequalities. I emphasize the role of labour market policy in incentivising higher education investments.

Second, I contribute to a body of literature on insurance-incentive trade-offs of disability policy by accounting for a broader set of behavioural responses to these programs. This paper fits among several studies that structurally model how disability policy drives labour market behaviour.⁹ Much of this literature

⁸For examples and surveys of related studies see Currie (2009), Case and Paxson (2010), Lundborg et al. (2014), ?, Almond et al. (2018), and Prinz et al. (2018).

⁹For instance, Gallipoli and Turner (2009), Bound et al. (2010), Kitao (2014), Low and Pistaferri (2015), Michaud and Wiczer (2018), Kostøl et al. (2019), and Kellogg (2021).

focuses on later onset disabilities, taking education level as given. However, it is crucial to account for the incentives SI has on early life decisions when considering the early-onset population because of the timing of their disability. To my knowledge, this is the first study to measure an insurance-incentive trade-off of disability policy with respect to educational investments.¹⁰ My results offer important insights into the design of DI and welfare programs when considering people affected by an early-onset disability.¹¹ I find that behavioural incentives matter for more than labour market decisions. Moreover, the moral hazard arising from these dynamic disincentives offers insight into the causes of application to disability programs.

Third, I contribute to the literature on the relationship between human capital investments, labour market conditions, and SI policies. Again, this paper aligns with studies linking education rates to the labour market environment.¹² The idea is that risks and public policies create incentives that distort behaviour in the labour market. If these distortions are large enough, they may also affect pre-entry decisions. Education is arguably the most important human capital investment decision before labour market entry. If the labour market distortions created by SA or DI are large enough, they can significantly impact the returns to schooling for this group. My research also relates to the literature studying how individuals make their education decisions given future uncertainty.¹³ My contribution is to evaluate the role of SI policy in partially insuring against uncertainty, affecting the expected value of self-insurance against future shocks through investing in post-secondary schooling.

The remainder of this paper is organized as follows. Section 2 outlines the Canadian policy environment. Section 3 presents the empirical model. Section 4 describes the dataset used for estimation. Section 5 discusses the estimation strategy and identification of the model's parameters. Section 6 reports the estimation results and evaluates the model's fit. Section 7 applies the model to conduct counterfactual policy experiments. Finally, Section 8 concludes.

2 Disability Policy Environment in Canada

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

This paper focuses on the two main programs providing long-term income assistance and replacement for

 $^{^{10}}$ Deshpande and Dizon-Ross (2023) conduct a similar study into dynamic disincentives of disability policy in the United States. They developed an experiment that provides parents of children with disabilities with information about the Social Security Income program, the US counterpart to the SA programs in Canada. Their focus was on how parental investments in their children were affected by this information. In contrast, I am interested in the post-secondary choices of individuals with early-onset disabilities.

¹¹This last point is relevant for theoretical literature on the design of SI policy, such as Golosov and Tsyvinski (2006)

 $^{^{12}}$ For instance, Flinn and Mullins (2015), Blundell et al. (2016), and Bobba et al. (2021).

¹³For instance, Carneiro et al. (2003), Cunha et al. (2005), and Navarro and Zhou (2017).

individuals affected by disability.¹⁴ These are the Canadian Pension Plan Disability (CPP-D), the federal DI program, and provincial SA programs, which offer means-tested welfare payments.¹⁵ The Canadian Pension Plan (CPP) is the federal retirement pension program that administers CPP-D. This section describes the main features of DI and SA in Canada.

2.1 Canadian Pension Plan Disability

DI in Canada delivers monthly financial transfers to applicants who are assessed and deemed eligible for the program. Eligibility requires applicants to be under the age of 65, not currently receiving Canadian Pension Plan (CPP) retirement benefits, to have made a predetermined number of contributions to CPP, and to be 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 expected to be indefinite or likely to result in death. CPP-D is a program for long-term disabilities and is not designed to insure against short-term disability spells. The severity of the disability is the applicant's ability to engage in "substantially gainful activity" in the labour market. That is, how productive a disabled individual is in a job they could be expected to hold given their qualifications relative to others doing the same work but who do not have a disability. Program adjudicators make a subjective assessment of an applicant's scope for substantially gainful activity given their disabiling condition and determinants of productivity. Adjudicators consider an individual's age, education, and work experience (Government of Canada, 2022).

The second eligibility requires that applicants have contributed to the CPP in four of the previous six years.¹⁶ Contributions to CPP are compulsory for working Canadians aged 18 to 70. Contributions equal a percentage of a worker's bounded employment earnings. In 2019, contributions equaled 4.95% of a worker's employment income up to \$55,900 in that year (Government of Canada, 2023). The contributions to CPP determine the monetary value, or generosity, of the CPP-D payments. The contributory period begins at age 18 and ends at age 65 or the year of death. It excludes years in which the applicant was receiving CPP-D benefits.

The monthly generosity of CPP-D is a function of an earnings index summarizing the average monthly earnings in the applicant's contributory period. In the calculation of the earnings index, applicants can drop certain months from their contributory period, which would reduce their final amount of CPP benefits.¹⁷ CPP-D payments are the sum of two components. The first component is equal to 75% of the applicant's potential CPP retirement benefits at the date of application. Potential CPP retirement benefits are equal to 25% of an earnings index that summarizes an applicant's bounded average earnings over their contributory period. The minimum bound to their earnings has been \$3,500 per year since 1996, and the maximum,

¹⁴I 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). I also do not model worker's compensation, which is only available to individuals injured at work and is not accessed by early-onset individuals in my data.

¹⁵In the following, I use CPP-D and DI interchangeably.

 $^{^{16}}$ Three of the previous six years if the applicant has contributed to the CPP for twenty-five years or more.

¹⁷First, each applicant is eligible to drop contributory months in which their children younger than seven years old. Second, applicants can drop a remaining percentage of their remaining contributory months with the lowest earnings.

which was \$53,600 in 2015, is updated yearly based on a measure of average wages. The second component is a deterministic flat-rate benefit indexed by the CPI each year.¹⁸

2.2 Provincial Social Assistance in Canada

Provincially administered SA programs are the main source of welfare transfers in the Canadian social safety net. These programs provide last-resort financial assistance to individuals facing barriers to their sustained employment and who have insufficient or volatile sources of income. SA is a form of last-resort social insurance accessible only to individuals who have exhausted all other means of assistance, including DI. SA programs do not impose a work requirement, extending eligibility to a broader population than DI.

SA is separately administered in each province. As such, the SA programs vary in eligibility criteria and the generosity of their welfare transfers by province. However, all SA programs share a similar overall structure (Employment and Social Development Canada, 2016).¹⁹ 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.

To determine eligibility, the means test calculates the net difference between an applicant's "assessed needs" and their income and assets. An applicant is eligible for SA if their assessed needs exceed the sum of their income and assets below an upper threshold. First, an applicant's "needs" may include variables like living expenses, family size and composition, and disability. Assessed income combines all earnings from market activities, such as paid employment or self-employment, with transfers from other government programs, such as DI.²⁰ Individuals may receive SA while earning from other sources, reducing the amount of benefits according to the program's replacement rate. SA can be revoked if sufficient effort is not taken on the beneficiary's part to receive other income support sources.

Recipients of SA typically receive monthly financial transfers composed of a basic assistance amount and, in some cases, a supplementary special assistance amount. The basic assistance amount is intended to cover essential living expenses, including food, shelter, and clothing. However, the cost of living can vary based on individual demographic characteristics, particularly disability status. Disabilities often entail higher living expenses and present additional employment barriers. Accordingly, all SA programs allocate additional resources for individuals with disabilities. Throughout the remainder of this paper, I refer to these disability-related benefits as SA-D.

¹⁸In 2018, the average CPP-D benefit received was just under \$1000 per month, half of which was the deterministic flat rate component (Employment and Social Development Canada, 2018).

¹⁹SA 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; Torjman and Makhoul, 2016).

 $^{^{20}}$ An applicant's financial assets include liquid assets, such as cash or convertible assets, and fixed assets, such as property. Exempt assets include those used for employment or transport, such as tools or automobiles, and assets related to savings plans used for education purposes, such as registered education savings plans. The combined fixed and liquid assets must not exceed a predetermined threshold, which varies by provincial jurisdiction. Additional details on SA programs can be found in Employment and Social Development Canada (2016) or Hillel et al. (2020).

3 Model of Education and Life-Cycle Labour Supply

I develop a life-cycle model of education investments and labour market decisions in an environment resembling the Canadian context. This model formalizes the relationship between net returns to education, the labour market, and SI policy. Using this model, I compare the relative importance of various factors contributing to the observed education gap between individuals with early-onset disabilities and those without. Additionally, the model enables analysis of the effects of policy reforms. For example, I predict behavioural responses to changes in the policy environment, comparing the incentive costs and insurance value of the counterfactual policies.

3.1 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). I adopt a parsimonious choice set for education. However, the relevant question is whether this level of parsimony is suitable for the objective of this paper, which is to measure the dynamic disincentives of labour market SI policy. Post-secondary education is defined as the completion of any degree beyond high school, representing a deliberate investment in human capital. By simplifying the educational choices, the model isolates the key trade-offs individuals face when considering further education under the influence of labour market SI policies. The education choice depends on a set of endowments that affect the expected return to each education level. Initial endowments include disability status, $d_0 \in \{0, 1\}$, which identifies the early-onset group. Individuals with $d_0 = 1$ have an early-onset disability and $d_0 = 0$ otherwise.

Education level, s, is chosen to maximize an individual's expected discounted lifetime utility. The expectation is a function of the labour market environment, accounting for the available SI policies and a set of risks that depend on s and d_0 . Those choosing s = 0 enter the labour market at age 19, and those choosing s = 1 enter at age 22. Time spent in the labour market lasts until, at most, age 65, after which everyone faces ten mandatory retirement periods and then dies.²¹ The lifespan of 75 years of age (T=57) is fixed for all individuals, and I assume there is no bequest motive. The life-cycle can be split into time in school, T^S , time in the labour market, T^L , and time in retirement, T^R .

Heterogeneity within disability

The impact of an early-onset disability can vary considerably, depending on, for instance, the severity or type of functional impairments. In the model, disability impacts individuals by reducing the returns to education and work and granting eligibility for SI, raising the expected value of not working.²²

The model considers a unitary notion of disability status but captures heterogeneity in the effects on education through two channels. Given d_0 , individuals receive an ability endowment, $a_i \sim iid \ Lognormal(\bar{a}^{d_0}, \sigma_{a^{d_0}}^2)$ and idiosyncratic shock to the cost of education, $\psi_i \sim iid \ N(\bar{\psi}^{d_0}, \sigma_{\psi}^2)$. The ability endowment determines

 $^{^{21}}$ The model assumes there is no mortality risk before the terminal period. In the data, mortality rates are trivially low before age 60 and, therefore, heavily discounted for the schooling decision. Moreover, differences in mortality by early disability status and education level only appear after age 60 and remain small in magnitude.

 $^{^{22}}$ In the data, there exists a negative correlation between disability and both employment and education. There also exists a positive correlation between disability and attachment to SI programs. The sign of these correlations is the same across all types and severity levels, although the magnitude varies.

earnings at labour market entry, representing unobserved factors influencing skill development before age 18. For instance, if certain types of disability interfere with classroom learning, if parents invest less in children with disabilities, or if tutoring expenditures are lower for children with disabilities, then this will impact the distribution of ability at age 18. The idiosyncratic cost of education accounts for unobserved factors -both pecuniary and non-pecuniary- that explain observed education choices that are inconsistent with ability sorting. For example, households with a child with a disability may face tighter budget constraints, making it challenging to cover the cost of tuition or housing during post-secondary education. Alternatively, it may be the case that certain aspects of university are more stressful with a disability.

3.2 Labour Market Environment

At each age $t \in T^L$, individuals choose whether to work, not work, or to apply for DI. Consequently, individuals can be in one of three labour market states: working, not working on SA, or not working on DI. These choices are made subject to uncertainty in future disability status, employment, and productivity, given the availability of partial insurance from SI policies.

Disability Risk

Disability status, $d_{it} \in \{0, 1\}$, evolves according to a first-order Markov process, where $d_{it} = 1$ if i is disabled in period t, and $d_{it} = 0$ otherwise. The transition probability for disability status is

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

$$\tag{1}$$

The risk of disability onset increases with age, t, and the likelihood of recovery decreases with age. The disability transition probabilities vary by d_0 , as early-onset disabilities represent a potentially different set of conditions that may evolve differently over the life-cycle. Disability risk is assumed to be exogenous to an individual's labour market choices, which is a standard assumption in the related literature (Low and Pistaferri, 2015; Michaud and Wiczer, 2018; Kostøl et al., 2019; Kellogg, 2021)

Search Frictions

While not working, individuals may enter employment if they receive an offer with age-dependent probability, $\lambda_t^{d_0,s}$. While employed, an individual may be exogenously displaced out of employment with age-dependent probability, $\delta_t^{d_0,s}$. Employed individuals may also choose to quit their jobs endogenously. These probabilities depend on s and d_0 to represent differences in search behaviour, institutional features, employer beliefs, and other barriers to working by early-disability status and education level (Acemoglu and Angrist, 2001; Dixon et al., 2003; Kitao, 2014; Morris et al., 2018).

Annual Earnings

An individual's potential earnings, W_{it} , are determined by a combination of potential work experience, P_{it} , current disability status, d_{it} , idiosyncratic shocks to productivity, $\epsilon_{it}^{d_0,s}$, and unobserved fixed heterogeneity,

 $v_i^{d_0,s}$. Potential earnings in period t are

$$\ln W_{it} = \mu_1^{d_0,s} P_{it} + \mu_2^{d_0,s} (P_{it}/100)^2 + \phi^{d_0} d_{it} + v_i^{d_0,s} + \epsilon_{it}^{d_0,s},$$
(2)
where $\epsilon_{it}^{d_0,s} = \epsilon_{it-1}^{d_0,s} + \xi_{it}^{d_0,s},$
 $\xi_{it}^{d_0,s} \sim iid \ N(0, \sigma_{\xi^{d_0,s}}^2) \text{ for } t > 0.$

The parameters governing potential earnings depend on initial disability status and education level. The second-order polynomial of experience provides curvature to the life path of potential earnings. The specificity of $\mu_1^{d_0,s}$ and $\mu_2^{d_0,s}$ to initial disability status lets d_0 affect the evolution of earnings over the life-cycle.²³ The return to potential experience also varies by education level, representing heterogeneity in the rate of productive skill accumulation on the job. The direct effect of a disability on productive human capital is captured by ϕ^{d_0} . This parameter captures a disability-induced loss of work-relevant skills, negatively affecting earnings.

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

An early-onset disability also impacts the development of productive skills during school. Unobserved fixed heterogeneity, $v_i^{d_0,s}$, is 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 assume the fixed effect is parameterized as

$$v_i^{d_0,s} = h^{d_0,s} a_i + \xi_{i0}.$$
(3)

The parameter $h^{d_0,s}$ scales an individual's ability differently depending on their initial disability status and chosen schooling level.²⁴ Initial earnings also depend on an initial shock, $\xi_{i0} \sim N(\bar{\xi}_0, \sigma_{\xi_0}^2)$, capturing earnings capacity that is unrelated to education and disability.²⁵

The Earnings Index

The earnings index summarizes an individual's earning history in T^L and determines the generosity of DI and retirement transfers.²⁶ The earnings index, e_t , is assumed to update each period given the previous period's earnings index, e_{it-1} , the individual's labour earnings in the current period, W_{it} , and age, t, according to

$$e_{it} = f(e_{it-1}, W_{it}, t) = \begin{cases} \frac{(t-1)e_{it-1}}{t} & \text{if } W_{it} < \underline{W} \\ \frac{(t-1)e_{it-1} + W_{it}}{t} & \text{if } W_{it} \in [\underline{W}, \ \bar{W}) \\ \frac{(t-1)e_{it-1} + \bar{W}}{t} & \text{if } W_{it} \ge \bar{W}, \end{cases}$$
(4)

 $^{^{23}}$ For instance, Cutler et al. (2006) study heterogeneity across education levels in one's ability to cope with a disability and its effect on the evolution of their life-cycle earnings.

 ²⁴This specification is similar to Flinn and Mullins (2015) and has the feature that human capital production technology is supermodular in ability.
 ²⁵For instance, this can represent earnings determinants such as networks from friends and family and non-human capital

²⁵For instance, this can represent earnings determinants such as networks from friends and family and non-human capital labour that can be supplied to the market. This term can also pick up regional differences in labour market policies, such as minimum wages.

 $^{^{26}}$ This index is similar to the average indexed monthly earnings measure that determines social security in the US.

where $e_{it} = 0$ for $t \in T^S$. The parameters \underline{W} and \overline{W} are the lower and upper bounds, respectively, on average earnings in period t. These are set to $\overline{W} = \$40,000$ and $\underline{W} = \$3,500$, which is consistent with the formula applied in the CPP and CPP-D programs.

Retirement

Individuals must exit the labour force and retire by age 65. However, individuals can choose to retire early, starting at age 60. Retirement income comes from an individual's pension benefits and old age security. Retirement benefits equal $0.25 * e_{it}$, which approximates the formula used in the CPP. Old age security is fixed at \$5,500, which approximates the average amount received from the Old Age Security Pension (OASP) program in Canada.²⁷ OASP helps supplement income for retirees with no CPP income, and individuals cannot receive OASP until age 65. If retiring early, an individual's retirement income is penalized 7.2% for each year 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.

Disability Insurance

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

To approximate the contribution requirement of CPP-D, I assume that individuals must have worked at least once to be eligible for DI. This requirement is captured by the binary variable ρ_{it} , which equals one if the contribution requirement is met and is zero otherwise.²⁸ I assume DI is only available for those with $d_{it} = 1.^{29}$ Hence, conditional on having applied to DI in the previous period, $m_{it-1} = 1$, the probability of acceptance is

$$PR(\mathbb{1}_{it}^{DI} = 1 | \rho_{it} = 1, d_{it} = 1, s_i) = \pi^s.$$
(5)

An individual's CPP retirement benefit is approximated as 25% of their earnings index, e_{it} . DI benefits equal 75% of their CPP retirement benefits plus a flat rate component.³⁰ The DI flat-rate component is set

²⁷This value is consistent with the average OASP income reported in the T1FF.

 $^{^{28}}$ This assumption has bite if individuals seeking DI with no work history work for only one period in order to meet the contribution requirement. However, this behaviour does not occur when estimating and simulating the model.

 $^{^{29}}$ As disability is measured based on limitations to daily activities, this assumption may miss some individuals with a health condition that automatically grants them access to DI. The sample of individuals who never report a disability but end up on DI in the data is trivial in the data and only occurs at the very end of the life-cycle.

³⁰I model DI generosity in a similar manner as Gallipoli and Turner (2009) and Milligan and Schirle (2019).

at \$4,365 and is known to agents in the model.³¹ Hence, DI generosity is given by

$$DI_t(e_{it}, b) = 0.1875 \ e_{it} + 4,365. \tag{6}$$

Individuals incur a utility cost of applying to DI, $C_{App}^{d_0,s}$. The application process can be lengthy and requires the applicant to submit a set of documents to prove they are eligible. The utility cost of this process may differ by schooling, as more educated individuals may be better equipped to complete the application process. Alternatively, education level may correlate with an individual's preference for self-sufficiency in the labour market, raising the application cost. I allow this disability cost to differ by d_0 , as early-onset individuals may be more familiar with the disability social safety net or have different preferences for selfsufficiency in the labour force.

Social Assistance

SA are means-tested benefits from anti-poverty programs. In reality, the value of benefits from SA differs by disability status and province of residence. In the model, I approximate provincial SA programs and the determination of SA benefits, which provide a lower bound to income for consumption. I assume the lower bound on consumption, $\bar{c}(d_{it}, pr_i)$, depends on disability status, representing the added SA-D resources for recipients affected by disability and province, pr_i . I assume there is 100% take-up of this program when not working or on DI. I define inc_{it} as an individual *i*'s income from all other sources. Then, the formula for SA is

$$SA(inc_{it}, d_{it}, pr_i) = \begin{cases} \bar{c}(d_{it}, pr_i) - inc_{it}, \text{ if } inc_{it} < \bar{c}(d_{it}, pr_i) \\ 0 & \text{otherwise.} \end{cases}$$
(7)

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

$$\bar{c}(d_{it}, pr_i) = \begin{cases} \bar{c}(0, pr_i) & \text{if } d_{it} = 0 \text{ or } \mathbb{1}^{SA-D} = 0\\ \bar{c}(1, pr_i) & \text{if } d_{it} = 1 \text{ and } \mathbb{1}^{SA-D} = 1, \end{cases}$$
(8)

where $\bar{c}(0, pr_i) < \bar{c}(1, pr_i)$.

Preferences

I assume a constant relative risk aversion (CRRA) utility function where consumption is non-separable from work and disability status.³² The utility functions for working ($L_{it} = 1$) and non-working ($L_{it} = 0$) individuals are given by

$$U(c_{it}, L_{it}; d_{it}) = \begin{cases} u^W(c_{it}; d_{it}) = \frac{(c_{it}e^{\theta d_{it} + \eta_1 + \eta_2 d_{i0}})^{1-\kappa}}{1-\kappa} & \text{if } L_{it} = 1\\ u^N(c_{it}; d_{it}) = \frac{(c_{it}e^{\theta d_{it}})^{1-\kappa}}{1-\kappa} & \text{if } L_{it} = 0. \end{cases}$$
(9)

 $^{^{31}}$ The real value of this amount has fluctuated between \$3,900 - \$4,500 over the calendar years spanned by the T1FF. The flat rate component in the model reflects a weighted average of this value over the years covered by my sample.

 $^{^{32}}$ Variants of this specification for preferences are common in related studies, such as, Low and Pistaferri (2015), Michaud and Wiczer (2018), and Kostøl et al. (2019).

This specification implies that disability and work may affect the marginal utility of consumption. I assume θ and η are negative, implying that workers or individuals with a disability require higher levels of consumption to have the same utility as non-working or non-disabled individuals.³³ These parameters capture the utility loss induced by work and disability, respectively. I include an additional utility cost to working with a disability, η_2 . The coefficient of risk aversion, κ , is assumed to be greater than one so that individuals are risk averse.

Individual's Problem in the Labour Market

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

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

s.t.
$$c_{it} = \tau (W_{it}L_{it}, DI_{it}) + SA(\tau (W_{it}L_{it}, DI_{it}), d_{it}) - F^d \mathbb{1}[t \ge 30, L_{it} = 1],$$
 (11)

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

Individuals decide to work or apply for DI in order to maximize their discounted lifetime utility, (10), subject to their budget constraint, (11), and the evolution of their earnings index, (12).³⁴ Utility from future periods is discounted by β . The expectation operator, \mathbb{E}_t , is conditional on the set Ω_t , which includes individual heterogeneity, $\{d_{i0}, a_i\}$, province of residence, pr_{it} , 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, ρ_{it} . I assume individuals are myopic in pr_{it} , so expect the same policy environment for all future periods. The agent's expectation is taken over all sources of risk, which include disability risk, idiosyncratic productivity risk, the job arrival rate, and the job destruction rate.

The budget constraint binds under the assumed preferences, implying income from all sources is consumed. W_{it} , DI, and SA are the monetary values of labour earnings, DI benefits, and SA benefits, respectively. Labour earnings and DI benefits are subject to income taxes through the function, $\tau()$, representing the Canadian combined provincial and federal tax system.³⁵ An individual receives labour income when employed, $L_{it} = 1$, and they receive DI benefits if they are eligible, $\rho_{it} = 1$, have chosen to apply, $m_{it} = 1$, and are accepted to the program. The monetary value of SA benefits is positive if the individual's income from other sources is below the poverty threshold, as described above. Last, individuals incur an additional monetary cost of working with a disability, F^d , when they are fifty years old and above (t > 30).

³³The utility cost of work nets out disutility from being on SA or DI.

 $^{^{34}}$ The model assumes income is consumed each period, as there is no market for savings. Justification for this assumption is that early-onset individuals have less scope for savings because of lower earnings. Moreover, this group has less incentive to save, given the higher generosity of the SI environment. As the research focus is education choices, the priority is to fit the earnings return to education. A related question is decomposing the value post-secondary into consumption or self-insurance, but this is outside the scope of this paper.

 $^{^{35}}$ The Canadian income tax system is a discrete set of tax rates and 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 details on the parameters of the tax and transfer system, refer to Section 6 of the Appendix.

Education Choice

The schooling decision is made at t = 0 based on the expected value associated with each schooling level, $V_0(d_0, a, s)$. The value functions, defined as in (10), depend on initial disability status, the ability endowment, and education level.³⁶ Individual *i* will choose to pursue post-secondary if

$$V_0(d_{i0}, a_i, s = 1) - V_0(d_{i0}, a_i, s = 0) - \psi_i \ge 0.$$
(13)

Education is a costly investment in terms of financial resources and utility and may be more costly in the presence of a disability. The model captures this by the idiosyncratic cost of post-secondary education, ψ_i . The cost associated with s = 0 is normalized to zero.

The inequality in equation (13) reflects how early disability may influence educational investments by affecting these value functions. With a continuum of rational, forward-looking agents, there is a group on the margin of choosing higher education. For SI policy, the expected future recipiency of SA or DI is contained within the value functions for each education level. Therefore, any changes in the expected recipiency necessarily shift the group of individuals on the margin.

3.3 Model Solution

I solve the model numerically via backward induction, as there is no analytical solution. The solution algorithm is conceptually straightforward: in each period, individuals make discrete choices (work, apply for DI, retire), and thus the policy functions are derived from conditional discrete choice problems. I begin by computing the terminal value at retirement (age 65) for all points in the state space. I then iterate backwards through time, using the value function approximated at t + 1 to determine optimal decisions at each age tgiven state variables $S_t = \{d_t, \epsilon_t, e_{t-1}, \rho_t\}$. In solving the model, I introduce additional i.i.d "taste" shocks to the utilities associated with each labour market state. Without these shocks, the likelihood function could exhibit sharp discontinuities over the parameter space, given the discrete nature of the problem.

Once I have solved for the individual's labour market decisions over the life cycle, I then solve the education choice policy function (at age 18) as a function of initial heterogeneity, $\{a, d_0, \psi\}$. This approach to solving the life-cycle model is standard in finite horizon discrete choice dynamic programming models. A detailed description of the solution algorithm and value functions are described in Section 3 of the Appendix.

4 Data: The Longitudinal and International Study of Adults

I estimate the empirical model using the Longitudinal and International Study of Adults (LISA) (Statistics Canada, 2018). LISA is a panel survey of over 11,000 Canadian households aged 15 and older. LISA consists of four biennial survey waves, starting in 2012, that cover a broad range of topics, including health, education, the labour market, social participation, and income. These data allow me to identify individuals with disabilities and the timing of onset. Moreover, LISA is supplemented with several administrative datasets. Most relevant are the T1 family files (T1FF), which contain rich disaggregated measures of personal income and transfer payments from annual income tax filings. Linking these datasets allows me to

³⁶The value of post-secondary, $V_{i0}(d_{i0}, a_i, s = 1)$, include the first three periods of utility during post-secondary and then labour market entry afterward.

build a comprehensive history of incomes and transfers between 1989 and 2017 for each observation in the sample. An advantage of these data is they are less likely to suffer from the measurement and coverage issues often associated with survey data. For instance, Meyer et al. (2009) show that survey measures of public transfers often suffer from respondents under-reporting, which can lead to overestimation of total income declines following the onset of disability.

Individual demographic information is obtained from the main survey waves of LISA. Each survey wave contains information about education, labour market status, changes in labour market status since the previous wave, job search activities, reasons for job loss, and details about limitations to daily activities, which I use to measure disability. Education is measured using a respondent's highest completed certificate. I flag individuals as having post-secondary education if they have completed any post-secondary certificate, including 2-year and 4-year college degrees, bachelor's degrees or higher, and vocational degrees. Individuals have a low education if their highest completed certificate is equivalent to high school or lower.

The 2014, 2016, and 2018 waves of LISA include measures of limitations to activities of daily living (LADL), which I use to identify disability.³⁷ The set of LADLs is derived from a short version of a set of disability screening questions (DSQ) developed by Statistics Canada for identifying individuals with disabilities in general population surveys (Grondin, 2016).³⁸ The activity limitations are self-reported in LISA. Respondents are asked a flow of categorical questions about the frequency of limitation for each LADL.³⁹ A respondent is flagged with a disability if reporting their condition to limit their activities "sometimes," "often," or "always." The age of disability on set is derived from a self-reported retrospective question, "at what age did you first start having difficulty or activity limitation?" Due to the retrospective nature of this question and the panel structure of the survey waves, there are instances where an observation reported different ages of onset. To address this, I use the minimum reported age of onset as the truth. An individual is flagged as having an early-onset disability ($d_0 = 1$) if the reported age of onset is eighteen or younger.

The T1FF tax records provide a panel of disaggregated measures of annual incomes and transfer payments from 1982 to 2017 for each respondent in the main survey waves of LISA. These data contain details on an individual's demographic characteristics relevant to their tax filings, such as age, sex, and province of residence. I measure annual earnings using paid employment income in the form of wages, salaries, and commissions (WSC), which are by far the largest component of market income. Annual transfers from CPP-D and SA are measured directly. I use these income measures to derive an individual's labour market status each year. An individual is considered employed if earning more than \$2,500 in WSC during that year. An individual is flagged as a DI beneficiary if reporting any positive income from CPP-D.⁴⁰ An individual is considered a non-participant receiving SA if not employed or on DI.

4.1 Sample Selection

My sample of interest consists of males with and without early-onset disabilities. The analysis is restricted to individuals aged 18 to 65 during the calendar years 1989 to 2017. Tax years prior to 1989 are excluded,

 $^{^{37}}$ The 2012 wave comprises only a small set of questions about disability and excludes information on the age of disability onset.

 $^{^{38}}$ Further details on the survey questions used to derive disability status are found in Section 2 of the Appendix

 $^{^{39}}$ Some cognitive conditions, such as developmental disability or learning conditions, are initially flagged based on diagnosis from medical professionals instead of the level of difficulty.

 $^{^{40}}$ There are very few instances in which a respondent reports positive WSC income and CPP-D income in the same year. In most such cases, the amount reported in WSC is below \$2,500.

as Canada Pension Plan Disability (CPP-D) benefits are not separately recorded from the Canada Pension Plan (CPP), and Social Assistance (SA) is not disaggregated from other non-taxable income in the available data before that year. Individuals with disabilities residing in institutions—defined as general hospitals, prisons, nursing homes, and specialized care facilities—are excluded from the sample,⁴¹ given that their severe conditions significantly constrain labour market participation. This study focuses on individuals with disabilities who are on the margin of pursuing post-secondary education. Additionally, individuals residing in the Canadian Territories are excluded. The final estimation sample includes approximately 650 individuals with early-onset disabilities and 8,000 individuals without, each observed for 20 periods on average.

4.2 Supplementary Data: Education Expenditure, Social Assistance and Taxes

The parameters of the SA policy in the model are determined from the annual report, "Welfare in Canada," produced by the Maytree Foundation (Maytree-Foundation, 2018).⁴² These reports calculate the maximum annual "total welfare income" a household may receive in each province and calendar year, which combines income from SA and provincial and federal tax credits. Maximum total welfare income is calculated for four distinct household types: single adult considered employable, single adult with a disability, single parent with one child aged two, and couple with two children aged 10 and 15. In calculating the total welfare income for each household type, it is assumed that the household received assistance for the entire year and had no earnings or assets, so they are entitled to the maximum amount of assistance; the household lived in the largest city in the respective province; the household lived in private market housing, and rent includes utility costs. I use the calculated maximum total welfare income for SA and SA-D in my model. The SA policy in the model will be slightly more generous than the policies in reality, as maximum total welfare includes additional provincial and federal tax credits. Section 6 of the Appendix describes how these data are mapped into policy parameters in the model.

The model's income tax system parameters are derived using data from the 2016-2 version of the Canadian Tax and Transfer Simulator (Milligan, 2016). This resource provides abundant information on federal and provincial income tax brackets and marginal tax rates, among other tax parameters, from 1962 to 2016. I combine federal and provincial rates to produce a distinct tax regime for each province in each calendar year covered by my study.

Lastly, I use the Canadian Tuition and Living Accommodation Costs Survey to calculate the level of consumption during post-secondary education (Statistics Canada, 2022). This survey collects data for fulltime students at publicly funded Canadian degree-granting institutions to provide information on tuition, additional compulsory fees, and living costs for an academic year. In the model, I set consumption during post-secondary education to be the same for all individuals. The rationale is that the idiosyncratic cost distribution will absorb differences in consumption during schooling across individuals. Consumption during post-secondary is calculated to be \$4450 per year, which is the weighted average price-adjusted tuition and ancillary fees from 1993 to 2018, where the weights are the sample density over these years.

⁴¹Institutional residents are individuals in general hospitals, prisons, nursing homes, and special care facilities for individuals with disabilities.

 $^{^{42}}$ These annual reports were formerly conducted by the National Council of Welfare until 2009.

5 Model Estimation and Identification

The model's parameters are recovered in two stages. The first stage calibrates a set of parameters to values from the related literature and estimates another set outside the life-cycle model's structure. The second stage estimates the remaining parameters using indirect inference, taking the parameters obtained in the first stage as given.

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 sufficiently risk-averse in the model. The discount factor is calibrated to $\beta = 0.9756$, the value used in Low and Pistaferri (2015).⁴³ The utility cost from disability, θ , is set to -0.488, as in Low and Pistaferri (2015). While θ is calibrated, the model accommodates additional flexibility in the effect of disability on preferences by estimating the disability-specific utility cost to working, η_2 .

Disability Risk



Figure 1: Disability Transition Probability Over the Life-Cycle

Notes: Life-cycle transition probabilities are derived with the 2012-2018 LISA survey waves. Transition probabilities are calculated for 5-year age bins, then smoothed with a LOWESS algorithm.

 $^{^{43}\}kappa$ is in a comparable range as estimated in Attanasio et al. (1999), Attanasio and Weber (1995), and Banks and Brugiavini (2001). In Low and Pistaferri (2015), β reflects the annual discount factor from Gourinchas and Parker (2002) and Cagetti (2003).

I estimate disability transition probabilities using the four survey waves of LISA, as disability status is not observed in the tax records.⁴⁴ The sample is partitioned by early-onset disability status and into nine distinct age groups. The age groups consist of eight five-year intervals beginning at age 25 and one category encompassing ages 18 to 25. Transition probabilities, $\gamma_{k,l}^{t,d_0}$, are estimated by calculating the likelihood of individuals transitioning into and out of disability status within each partition. The estimated age-specific transition probabilities across the life course are presented in Figure 1.

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

Social Assistance

The SA thresholds, $\bar{c}(d_{it}, pr_i)$, are derived using the maximum total welfare incomes calculated in the Maytree annual reports. These reports provide maximum benefit levels for individuals with and without disabilities across each provincial jurisdiction and calendar year. I represent the SA policy environment using a two-element "couplet," specifying the maximum benefit available under standard SA and disability-targeted SA (SA-D). With data spanning 10 provinces over 29 years, this yields 290 distinct couplets. Accommodating such a large number of unique SA regimes in the model is computationally infeasible. To address this, I use the Hartigan–Wong k-means clustering algorithm to group "similar" couplets based on Euclidean distances (Hartigan and Wong, 1979). The couplets are consolidated into two clusters, $pr_i \in \{1, 2\}$. The choice of two regimes ensures that each regime has sufficient support within the data while providing identifying variation to the policy environment. The resulting SA thresholds used in (8) are,

$$\bar{c}(d_{it}, pr_i) = \begin{cases} 5,300 & \text{if } d_{it} = 0 \text{ and } pr_i = 1\\ 8,378 & \text{if } d_{it} = 1, \ \mathbbm{1}^{SA-D} = 1, \ \text{and } pr_i = 1\\ 7,406 & \text{if } d_{it} = 0 \text{ and } pr_i = 2\\ 10,261 & \text{if } d_{it} = 1, \ \mathbbm{1}^{SA-D} = 1, \ \text{and } pr_i = 2 \end{cases}$$

Accordingly, one cluster corresponds to provinces with less generous SA policies, while the other reflects more generous welfare provisions. Further methodological details regarding the clustering process are provided in Section 6 of the Appendix.

Employment Risk

The job arrival and destruction rates, $\lambda_t^{d_0s}$ and $\delta_t^{d_0s}$, are estimated directly from the main survey waves of LISA. To estimate the job destruction rate, I consider a sub-sample of individuals who reported being employed in the previous survey wave. I create a dummy variable equaling one if this individual reported being fired or laid off from this previous employment and zero otherwise. This dummy is regressed on age, and the resulting estimates are used to predict an age-specific job loss probability. Estimations are conducted

 $^{^{44}}$ This implicitly assumes the transition probabilities are the same in 2012-2018 as in the calendar years covered in the tax records.

separately by early disability status and educational level.

For the job arrival rate, I utilize a retrospective variable measuring the monthly labour market status over the 36 months preceding the interview. The data are reshaped into a monthly panel, and I generate a dummy variable equal to one if the respondent is employed (part-time or full-time) that month. I created another dummy variable that equals one if the respondent reported an active job search in the preceding month. The employment dummy is then regressed on the lagged-search indicator interacted with age to estimate job arrival probabilities by age. It is important to note that this approach captures only the probability of transitioning from job search to employment. It does not reflect the frequency of job offers or the rate at which offers are declined. Nevertheless, this reduced-form transition probability introduces labour market frictions into the model, influencing the value of SI policies relative to employment. Additional details on the computation of job arrival and destruction rates are provided in Section 7 of the Appendix.

5.1 Indirect Inference

Estimation of the remaining structural parameters is achieved by 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. The main ingredient of indirect inference is an auxiliary model that is made up of moments in the data providing identifying information for the remaining structural parameters. Indirect inference chooses the economic model's parameters such that the auxiliary model estimated using the observed data is as close as possible to the auxiliary model estimated using data simulated from the economic model. The observed data is an unbalanced panel, and I replicate censoring in the observed data when calculating the moments from the simulated data.⁴⁵

The set of parameters estimated via indirect inference, denoted Θ , includes the parameters governing the earnings process, the distribution of initial heterogeneity, probabilities of acceptance to DI and SA-D, and the utility costs associated with work and DI application. The estimated parameters, $\hat{\Theta}$, are obtained by minimizing the weighted sum of squared deviations between observed and simulated moments, as specified by:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \bigg\{ \sum_{k=1}^{K} \big[(M_{kN}^d - M_{kr}^m(\Theta))^2 / Var(M_{kN}^d) \big] \bigg\},$$
(14)

where the sum is over K moments, M_{kN}^d is the kth empirical moment estimated over N observations, and $M_{kr}^m(\theta)$ represents the corresponding simulated moment computed at parameter vector Θ across r replications. Each squared deviation is weighted by the variance of the corresponding empirical moment, $Var(M_{kN}^d)$.⁴⁶ An exception is made for moments of the rate of post-secondary, for which I use an order of magnitude smaller than the variance of the data moment.⁴⁷ The auxiliary model incorporates 217 moment conditions to estimate 34 structural parameters. I simulate the life-cycle decisions of 15,000 indi-

 $^{^{45}}$ A further desription of the censoring procedure is found in Section 4 of the Appendix.

⁴⁶I follow Blundell et al. (2016) in using $Var(M_{kN}^d)$ as the weighting matrix to avoid potential small-sample biases associated with the asymptotically optimal weighting matrix, as discussed in Altonji and Segal (1996).

⁴⁷This adjustment ensures that the model can match the observed schooling distribution, which otherwise receives relatively lower weight compared to more precisely estimated life-cycle earnings and employment moments.

viduals—three replications of 5,000 individuals each.⁴⁸ Tables presenting the moment estimates from the data and model are provided in Section 9 of the Appendix.

5.1.1 Earnings Parameters

The parameters of the earnings process are identified using life-cycle variations in annual employment income, disability status, and education level. The primary set of identifying moments are derived from estimating an earnings process related to (2). The earning parameters, $\hat{\mu}_1^{d_0,s}$, $\hat{\mu}_2^{d_0,s}$, and ϕ , are identified by matching the coefficients from the following first-difference regression model,

$$\Delta \ln W_{it} = \hat{\mu}_1^{d_0,s} \Delta P_{it} + \hat{\mu}_2^{d_0,s} \Delta P_{it}^2 / 100 + \hat{\phi} \Delta d_{it}^* + \epsilon_{it},$$
(15)

where ΔX_{it} denotes the first difference of variable X_{it} for individual *i*, ϵ_{it} is a regression error, and $d_{it}^* = 1$ for all periods following the initial onset of disability, and $d_{it}^* = 0$ otherwise. Transitions in and out of disability status are not observed in the income tax records, only the year of initial onset. However, as disability transition probabilities are estimates outside of the model and are taken as given, the first onset helps pin down ϕ . It is important to note that (15) does not account for potential selection into employment. However, this limitation is addressed within the structural model, which explicitly simulates employment decisions. As such, the estimation procedure selects structural parameters to replicate the selection patterns observed in the empirical data, thereby implicitly correcting for this source of bias.

Given $\hat{\mu}_1^{d_0,s}$, $\hat{\mu}_2^{d_0,s}$, and $\hat{\phi}$, I can calculate the residual from (15),

$$\hat{\nu}_{it}^{d_0,s} = \Delta \ln W_{it} - \hat{\mu}_1^{d_0,s} \Delta P_{it} + \hat{\mu}_2^{d_0,s} \Delta P_{it}^2 / 100 + \hat{\phi} \Delta d_{it}^*.$$
(16)

The residual, \hat{v}_{it} , is a similar object as the idiosyncratic productivity shocks from (2), given that $\epsilon_{it}^{d_0,s}$ is assumed to follow a random walk.⁴⁹ I use the sample variance of $\hat{\nu}_{it}^{d_0,s}$ to pin down the variance of the productivity shock, $\sigma_{\ell^{d_{0,s}}}^2$.⁵⁰ Additional moments that I use to help pin down the life-cycle pattern in earnings and earnings penalty of a disability include estimates from regressing log earnings on a quadratic in age and indicator for first disability onset. Moreover, I match five earnings quantiles conditional on early disability status and school level to fit the aggregate distribution of earnings over all years within each group. I also compute the mean and variance of annual earnings in the first three periods in the labour market to help pin down the distribution of initial earnings.

Policy Parameters

To identify the probability of acceptance to DI, π^s , I use rates of DI receipt and flows onto DI. While DI applications are unobserved in the data, the model simulates DI applications for a given guess of structural parameters. The resulting moments reflect the simulated decision to apply for DI. I match estimates from

⁴⁸In each replication, half of the sample is assigned early-onset disability status, diverging from the empirical distribution. However, since all moments are stratified by education level (s) and early-onset disability status (d_0) , and no compositional moments across d_0 are used, this does not bias the estimation. ⁴⁹To see this, from (2), we can write $\xi_{it}^{d_0,s} = \Delta \epsilon_{it}^{s,d_0} = \Delta \ln W_{it} - \mu_1^{d_0,s} \Delta E_{it} + \mu_2^{d_0,s} \Delta E_{it}^2/100 + \phi \Delta d_{it}$. ⁵⁰Earnings variation related to changing disability status will be contained in $\hat{\xi}_{it}^{d_0,s}$, as d^* is absorbing. However, the model

will replicate this latency as disability status simulated using predetermined parameters.

the following two linear regressions:

$$\mathbb{1}(DI_{it}) = \beta_0^{s,d_0} + \beta_1^{s,d_0}t + \beta_2^{s,d_0}t^2 + \beta_3^{s,d_0}t^3, \text{ and}$$
$$\mathbb{1}(DI_{it} = 1 \& DI_{it-1} = 0) = \beta_4^{s,d_0} + \beta_5^{s,d_0}t + \beta_6^{s,d_0}t^2 + \beta_7^{s,d_0}t^3,$$
(17)

where $\mathbb{1}(DI_{it})$ is an indicator variable equalling one if individual *i* is attached to DI in period t and $\mathbb{1}(DI_{it} = 1 \& DI_{it-1} = 0)$ is an indicator variable equalling one if individual *i* flowed onto DI in period t.⁵¹ These moments help the model fit life-cycle trends in DI application and enrollment, which are mostly zero in early life and then grow at an increasing rate after age 45 for all groups. Application costs to DI are identified using the average non-participation rate and the earnings distribution of applicants in the t-7 to t-2 years prior to enrollment.⁵² The intuition is that if the distribution of earnings is simulated to be too lower on average than the data, a larger guess for application costs will simulate DI enrollment from individuals with a relatively higher return to DI. These are individuals with higher pre-application earnings, as the value of DI is dependent on the earnings index.

The acceptance rate of SA-D, π^{SA-D} , is identified using the estimated change in the probability of employment following the onset of disability for the not-early-disabled sample. Prior to disability onset, the work decision is made based on the comparison between labour income and non-employment income, which is determined exogenously by the SA policy parameters. Following disability onset, the expected income from non-employment depends on the likelihood of receiving the more generous SA-D benefits, governed by π^{SA-D} . The parameters governing labour income are identified separately. Hence, the post-onset employment rate is a monotonically decreasing function of π^{SA-D} , which will calibrate to match the drop in employment after disability onset.

Preference Costs to Work

The preference parameters to be estimated include the utility cost of work, η_1 , an additional utility cost from working with a disability, η_2 , and the monetary cost of working with a disability at older ages, F^d . These parameters are identified through variations in employment patterns across age, disability status, and SA regime. For instance, η_1 governs a general disutility from working and is identified by matching observed employment rates across the life cycle. As η_1 increases, the model predicts lower labour supply, holding constant the consumption gains from employment—which are separately determined by the earnings process, policy settings, and taxation. Variation in employment rates across SA regimes helps to refine the estimate of η_1 . The additional disutility from working with a disability, η_2 , is identified by differences in employment trajectories over the life cycle between individuals with and without early-onset disabilities and how these patterns differ across SA regimes. Finally, the parameter F^d , capturing the monetary penalty of working with a disability in older age, is identified by targeting employment rates specifically at older ages among disabled individuals.

 $^{^{51}}$ 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 that disability severity and education level.

 $^{^{52}}$ It is assumed that individuals must apply to receive DI in t from unemployment in t-1.

5.1.2 Ability and Return to School

The ability distribution and return to post-secondary, $\{h_1^0, h_1^1, \bar{a}_0, \bar{a}_1, \sigma_{a^0}^2, \sigma_{a^0}^2, \bar{\xi}_0, \sigma_{\xi_0}^2\}$, are identified from the distribution of individual fixed effects that are calculated using estimates from (15). Given $\hat{\mu}_1^{d_0,s}, \hat{\mu}_2^{d_0,s}$, and $\hat{\phi}$, I to compute

$$\hat{v}^{d_0,s} = (T_i^w)^{-1} \sum_{t}^{T_i^w} \left(ln W_{it} - \hat{\mu}_1^{d_0,s} P_{it} - \hat{\mu}_2^{d_0,s} P_{it}^2 / 100 - \hat{\phi} d_{it}^* \right), \tag{18}$$

where T_i^w is the number of years the individual i has been observed working. Recall that $v^{d_0}(a_i, s_i) = h_s^{d_0}a_i + \xi_0$, so the distribution of $\hat{v}^{d_0,s}$ maps to the distribution's of a_i and ξ_0 . I match the mean, variance, and cutoff values of the 10th, 25th, 50th, 75th, and 90th quantiles of the distribution of $\hat{v}^{d_0,s}$ separately by s and d_0 (28 moments). Moments for those with s=0 pin down the selected ability distribution given the normalization that $h_0^0 = h_0^1 = 1$. Moments for those with s=1 pin down the selected ability distribution scaled by $h_{d_0}^1$. The distribution of ξ_0 is independent of s and d_0 , and ξ_{i0} is essentially noise as it is revealed after the schooling decision is made and doesn't affect selection. Hence, its distribution is pinned down to jointly fit the distributions of $\hat{v}^{0,0}$, $\hat{v}^{0,1}$, $\hat{v}^{1,0}$, and $\hat{v}^{1,1}$.

5.1.3 Parameters of Psychic Cost to School

The education decision is made conditional on initial endowments, which include disability status (d_0) , ability (a), SA regime (pr), and the idiosyncratic cost to education (ψ^{d_0}) . The psychic cost parameters are identified using a linear probability model of education choice. Conditional on d_0 , I regress s on $\hat{v}^{d_0,s}$, pr_i , and their interaction. The psychic cost variance, $\sigma_{\psi^{d_0}}$, is identified by the variance of the residuals from these models.

6 Estimation Results

The remaining sections review the estimation and implications of the structural model. First, I discuss the reasonableness of the estimated parameters and the fit of the model to key moment counterparts in the data. I then use the estimated model to investigate the most important factors contributing to the gap in educational attainment between individuals with and without an early-onset disability. Lastly, I use the model to conduct counterfactual experiments to predict how early-onset individuals respond to changes in policy.

Table 1 reports the indirect inference estimates for parameters characterizing the distributions of latent initial heterogeneity. First, the mean and variance parameters for the ability distribution and idiosyncratic cost are separately reported for individuals with (right) and without an early-onset disability (left).

Individuals with early-onset disabilities exhibit lower mean ability levels, translating into approximately \$1,850 less in initial earnings at labour market entry compared to those without early disabilities. Moreover, 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

$d_0 = 0$		$\underline{d_0} =$	= 1
Parameter	Estimate	Parameter	Estimate
$ar{a}^0 \ \sigma^2_{a^0} \ ar{\psi}^0$	$1.9694 \\ (0.012) \\ 0.0000 \\ (0.001) \\ 0.0084 \\ (0.000)$	$ar{a}^1$ $\sigma^2_{a^1}$ $ar{\psi}^1$	$1.9370 \\ (0.003) \\ 0.0044 \\ (0.001) \\ 0.0469 \\ (0.009)$
$\sigma^2_{\psi^0}$	(0.0073)	$\sigma^2_{\psi^1}$	(0.0102) (0.001)
$rac{ ext{Initial Earn}}{ar{\xi_0}} \ \sigma^2_{\xi_0}$	ings Shock 2.0076 (0.000) 0.0081 (0.001)		(0.001)

Table 1: Estimates of Parameters for Individual Heterogeneity

for early-onset individuals.⁵³ Individuals with an early-onset disability incur twice the idiosyncratic cost of post-secondary education on average, and this cost is five times as volatile. The utility cost equates to an average reduction in yearly consumption of \$578 for not early disabled individuals and \$1258 for early-onset. The mean of ξ_0 , presented in the bottom panel, implies that nearly half of initial log earnings is unrelated to education.

Table 2 presents estimates of the annual earnings process, again separately reported for individuals with and without an early-onset disability on the right and left panel, respectively. The direct effect of a disability, which is the same by d_0 , results in a 9% reduction in annual earnings. Skill accumulation during post-secondary scales initial ability by 3.1% for individuals with early-onset disabilities, compared to 4.2% for those without such disabilities. This difference, combined with the difference in mean endowed ability, translates into an approximate \$5,450 earnings gap at labour market entry for individuals who complete post-secondary education. This result is consistent with disabilities disrupting the efficiency of human capital accumulation during post-secondary schooling.

Earnings growth over the life cycle is determined by the parameters μ_1^{s,d_0} and μ_2^{s,d_0} . The estimates reveal that individuals with early-onset disabilities who do not pursue post-secondary education (s = 0) experience stagnant earnings growth throughout their working lives. In contrast, those with early-onset disabilities who attain higher education see significantly stronger earnings growth, indicating that the returns to post-

Notes: Table presents indirect inference estimates of the structural parameters. Standard errors are in parenthesis below point estimates. Standard errors are computed using the formula for the asymptotic variance, corrected for simulation error, provided in Gourieroux et al. (1993).

 $^{^{53}}$ 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.

$d_0 = 0$		$d_0 = 1$		
Parameter	Estimate	Parameter	Estimate	
ϕ_0 h^1	-0.094 (0.005) 1.040	$\phi_1 \ h^0$	-0.094 (0.005) 1.031	
$\mu_1^{0,0}$	(0.004) 0.110 (0.002)	$\mu_1^{0,1}$	(0.002) 0.074 (0.003)	
$\mu_2^{0,0} \ \sigma_{\xi^{0,0}}^2$	-0.216 (0.009) 0.016 (0.002)	$\mu_2^{0,1}$ $\sigma_{\xi^{0,1}}^2$	-0.140 (0.014) 0.018 (0.004)	
$\mu_1^{1,0}$	(0.002) 0.128 (0.003) 0.252	$\mu_1^{1,1}$	$(0.004) \\ 0.143 \\ (0.004) \\ 0.200$	
μ_2 $\sigma^2_{\xi^{1,0}}$	$\begin{array}{c} -0.252 \\ (0.008) \\ 0.014 \\ (0.001) \end{array}$	$\mu_1 \\ \sigma^2_{\xi^{1,1}}$	$(0.299) \\ (0.011) \\ 0.010 \\ (0.001)$	

Table 2: Estimates of Parameters for Earnings Process

Notes: Table presents indirect inference estimates of the structural parameters. Standard errors are in parenthesis below point estimates. Standard errors are computed using the formula for the asymptotic variance, corrected for simulation error, provided in Gourieroux et al. (1993).

secondary education are particularly high for this group. For individuals without early-onset disabilities, earnings growth is more uniform across education levels. These dynamics are illustrated more clearly in the life-cycle earnings profiles presented below. Lastly, earnings are the most volatile for early-onset individuals with low education, and the variance of productivity shocks increases with education regardless of early disability status.

The remaining parameter estimates, which relate to preferences and the policy environment, are reported in Table 3. The utility cost of working, η , equates to approximately 8% of annual consumption for all individuals. This cost increases by 9% for those with an early-onset disability. The added cost of working with a disability at old ages equal approximately \$2,500, which is in a similar ballpark as the fixed costs estimated in Low and Pistaferri (2015) for individuals with moderate disabilities.

The likelihood that individuals with disabilities are awarded SA-D benefits is 83.4%, reflecting a less stringent eligibility threshold compared to DI. As for DI, acceptance probabilities are modestly lower among applicants with post-secondary education. Notably, the model's unconditional DI acceptance rate of 46.3% closely aligns with the actual unconditional acceptance rate of 43% reported for the 2014–2015 fiscal year (Office of the Auditor General of Canada, 2015). The acceptance rate is not targeted in estimation and serves as external validation of the model's DI program approximating CPP-D. Finally, DI applications impose a utility cost on applicants, with the magnitude of this cost differing by early-disability status.

Utility Parameters		Policy Parameters		
Parameter	Estimate	Parameter	Estimate	
η_1 η_2 F^d	$\begin{array}{c} -0.0761 \\ (0.032) \\ -0.0948 \\ (0.055) \\ 2504.079 \\ (152.565) \end{array}$	π^{SA} π_0 π_1 $C^{0,0}_{app}$ $C^{1,0}_{app}$ $C^{1,1}_{app}$	$\begin{array}{c} 0.8239\\ (0.072)\\ 0.4734\\ (0.053)\\ 0.4525\\ (0.013)\\ -0.0015\\ (0.000)\\ 0.0011\\ (0.000)\\ 0.0011\\ (0.000)\\ \end{array}$	

Table 3: Estimates of Parameters for Utility and Earnings Process.

Notes: Table presents indirect inference estimates of the structural parameters. Standard errors are in parenthesis below point estimates. Standard errors are computed using the formula for the asymptotic variance, corrected for simulation error, provided in Gourieroux et al. (1993).

6.1 Model Fit

	Moment	Simulated
Early-Onset	0.460	0.457
Not Early Disabled	(0.037) 0.640 (0.012)	0.638

Table 4: Rate of Post-Secondary in Data and Simulated from Model

Notes: Standard errors of data moment in parenthesis below.

Next, I compare the fit of the estimated model relative to its moment counterparts in the data. As it can be difficult to interpret the values of estimated parameters in a large structural model, contrasting the true moments with moments calculated using data simulated from the model helps to validate the estimated parameters. Of first order, Table 4 reports 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, 18 percentage points, is identical to the gap observed in the data.



Figure 2: Model Fit: Life-cycle Annual Earnings

Note: Figure Plots predicted annual earnings in the actual data and simulated data. Predicted annual earnings are conditional on age, education level, and early-onset disability status, in the actual data and simulated data. Estimates used to generate predicted annual earnings are in **Table X** of Section 9 in the Appendix.

The main benefit of post-secondary education is the return to labour market earnings. Figure 2 presents the predicted life-cycle annual earnings profile separately by early disability status and education level. The left figure pertains to the not early disabled group, where the dotted lines are from the actual data, and the more solid lines are from data simulated with the model. The model recovers a similar age-earnings profile as in the data. Moreover, the it recovers a similar earnings premium to post-secondary schooling. The right figure pertains to the early-onset group. Again, the model recovers a very similar life-cycle earnings profile for each education group.

Annual earnings represent the return to working relative to the outside option of not working and collecting government transfers. Figure 3 reports aggregate employment rates separately by early disability status, education level, and age groupings. The darker bars are conditional on post-secondary education, and the lighter bars are conditional on low education. The right side graph shows rates simulated from the model, and the left side graph shows rates from the actual data. First, the model reproduces the increase in employment with age among individuals without early-onset disabilities, shown in the top figure, within each education level. Additionally, the model reproduces a higher employment rate by education very well. The lower figure reports the same for the early-onset group. Again, the model does a good job in reproducing the gap in aggregate employment by age groups and the positive relationship between education and employment.

Lastly, Figure 4 presents life-cycle rates of DI by early disability status and education level over all years in the labour market. The estimated model matches the life-cycle profile in the rate of DI, with a rapid flows onto the program as individuals near retirement. The model under predicts the average enrollment for all individuals. However, few individuals transition onto DI in the data, and the moments related to DI rates are very imprecise. Consistent with the data, the model predicts that early-onset individuals have the



Figure 3: Model Fit: Aggregate Employment

Note: Figure plots conditional employment rates in the actual data and simulated data. Employment rates are conditional on education, early-disability status, and age group, where "Young" is less than 45 years of age, and "Old" is 45 years of age or more. Moments used to generate the figure are located in **Table X** of Section 9 in the Appendix.



Figure 4: Model Fit: DI Rate by Age

Note: Figure Plots predicted rate of DI in the actual data and simulated data. DI rates are conditional on age, education level, and early-onset disability status, in the actual data and simulated data. Estimates used to generate predicted annual earnings are in **Table X** of Section 9 in the Appendix.

highest percentage of their population on DI.

6.2 Decomposing the Education Gap

The model predicts an 18 percentage-point gap in post-secondary attainment between individuals with and without an early-onset disability, hereafter referred to as the *baseline gap*. This section uses the estimated model to decompose the baseline gap and isolate the role of SI policy relative to the other model elements. To do so, the model's structural parameters have been grouped into several contributing factors, each consisting of parameters that differ by d_0 . The predicted education gap is a function of these factors, as differences in the underlying parameters by d_0 generate differences in the net returns to post-secondary. The contribution of each factor is determined by "shutting off" that factor in a specified sequence, where a contributing factor is shut off by setting the relevant parameters equal to those for individuals without an early-onset disability ($d_0 = 0$). For instance, if idiosyncratic costs are a contributing factor, then this factor is shut off when setting $\bar{\psi}^1 = \bar{\psi}^0$ and $\sigma_{\psi^1}^2 = \sigma_{\psi^0}^2$. I specify seven primary contributing factors: labour market policy, idiosyncratic costs, endowed ability and return to school, preferences, earnings growth, labour market uncertainty, and disability risk.

I perform two separate decompositions to assess the relative importance of contributing factors to the baseline gap. First, I calculate the ceteris-paribus marginal contribution of each factor by shutting it off individually while holding all other factors at their baseline value. I then re-simulate the model and calculate the change in the education gap. This approach reveals the relative importance of each factor in the baseline economy. However, these marginal contributions will not generally sum to the entire baseline gap nor account for the interactions that arise when multiple factors are shut off simultaneously.

Second, I apply a Shapley decomposition to measure the expected marginal contribution of each factor to the overall education gap (Shorrocks et al., 2013). This method considers every possible order in which factors could be shut off and calculates each factor's marginal contribution to the gap when it is removed. The expected marginal contribution of a factor is calculated as the average of these contributions across all possible sequences, capturing what that factor's expected contribution would be as if the order of removal were chosen randomly. Unlike simpler decomposition methods, the Shapley approach guarantees each factor is assigned a unique, permutation-invariant contribution and ensures these factor-level contributions add up exactly to the total gap. The detailed steps for the Shapley decomposition are provided in Section 8 of the Appendix.

Table 5 presents the results of each decomposition method. The top two rows show results from the ceteris paribus decomposition, and the bottom two rows show results from the Shapley decomposition. In each case, the top row presents each factor's contribution to the baseline gap and the contribution as a percentage of the baseline gap underneath.

In both decompositions, idiosyncratic costs emerge as the primary contributor to the baseline education gap. Individual-specific costs—encompassing both psychic and pecuniary barriers to education—significantly discourage the educational attainment of early-onset individuals. Similarly, the estimated differences in endowed ability and in the returns to post-secondary education also strongly contribute to the observed gap. In the first decomposition, the marginal effect of equalizing the ability distribution and return to schooling cut the gap in half.

A second notable finding is that the education gap widens when earnings growth is "equalized" between

	Idiosyncratic Cost	Ability and RTS	Earnings Growth	Labour Market Uncertainty	Preferences	Labour Market Policy	Disability Risk
Ceteris Paribus Decompositio	on						
Contribution to Gap	0.160	0.091	-0.157	0.031	0.025	0.077	-0.010
% Change to Baseline Gap	0.889	0.503	-0.871	0.172	0.140	0.429	-0.053
Shapley Decomposition							
Contribution to Gap	0.162	0.076	-0.177	0.034	0.059	0.034	-0.009
% Change to Baseline Gap	0.900	0.421	-0.985	0.188	0.326	0.191	-0.050

Table 5: Decomposition of the Baseline Education Gap

Notes: Several contributing factors encompass multiple sub-factors. Labour market policy combines SA-D and DI, endowed ability and RTS combine ability distributional parameters and h^{d_0} , labour market uncertainty combines job loss rate, job offer rate, and productivity risk, and preferences combine utility cost of disability, utility cost of working with a disability, and the cost of working with a disability in old age.

individuals with and without an early-onset disability. Early-onset individuals gain a notably large boost in their earnings-growth parameters $(\mu_1^{d_0,s}, \mu_2^{d_0,s})$ when they complete post-secondary education, in contrast to an earning profile that remains more stagnant and peaks at a lower level for those with low-education. By comparison, non-early-disabled individuals have relatively more similar life-cycle earnings profiles across education levels. Thus, imposing the same earnings-growth rates across both groups diminishes the advantage post-secondary education offers early-onset individuals, reducing their incentive to pursue higher education and ultimately expanding the education gap.

The calculated contribution of labour market policy differs in the two decompositions. The contribution of DI and SA-D in the ceteris paribus decomposition decreases the baseline gap by 42.9%. This indicates that disability-related policy is a key driver of the gap under baseline parameter values. However, the expected marginal effect of this policy reduces the baseline gap by only 19.1%. Once other contributing factors are neutralized, the marginal role of labour market policy in shaping education decisions diminishes. The reason is that, for the most part, shutting off the other factors encourage post-secondary education. Consequently, fewer individuals remain on the margin where the net utility of education is near zero, making changes in disability-related benefits (such as SA-D) less important in education choices. When averaged across these various counterfactual conditions, the marginal impact of labour market policy is, therefore, smaller. Despite this reduced marginal effect, labour market policy still plays a significant role in disincentivising post-secondary investment. These findings suggest that addressing other sources of the baseline education gap may also help reduce the dynamic disincentives embedded in disability policy.

Several contributing factors encompass multiple sub-factors. Labour market policy, for instance, includes both SA-D and DI policies. Endowed ability and returns to school reflect differences in the distribution parameters of ability endowments and educational returns. Labour market uncertainty captures the combined effects of labour market frictions and productivity risk. Preferences account for the utility cost of disability, the cost of working with a disability, and the cost of disability in old age. Table 6 presents the results of a secondary decomposition, which recovers the contribution of each sub-factor to the respective primary factor as described in (Shorrocks et al., 2013). The first column shows the proportion contribution of each sub-factor to its respective primary factor, while the second column displays the proportion contribution to the overall gap.

	Proportion of Primary Factor	Proportion of Baseline Gap
Ability and RTS $ar{a}^{d_0}, \sigma^2_{a^{d_0}}, \sigma^2_{a^{d_0}}$	0.510 0.490	0.215 0.207
Earnings Growth $\mu_1^{s,d_0}, \mu_2^{s,d_0}$	0.972	-1.029
φ Labour Market Uncertainty $λ_{t_{0}}^{d_{0},s}, \delta_{t}^{d_{0},s}$	0.028	0.178
$\sigma_{\psi^{d_0}}^2$ $Preferences$ η_2	0.050 0.001	0.009
θ, F ^d Labour Market Policy	0.999	0.326
DI	0.008	0.002

Table 6: Contribution of Secondary Factors in Shapely Decomposition

Within the policy, SA-D is the main contributor to the education gap. This program, which raises the outside option of work, is particularly valuable for early-onset individuals who face greater adversity in the labour market. The DI program has minor effects on education for two main reasons. First, individuals generally flow onto DI in the second half of their 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.

7 Counterfactual Policy Experiments

A main advantage of formalizing and estimating a structural model is the ability to evaluate the effects on individual behaviour, welfare, and the government budget from changes to the policy environment. The decomposition exercise reveals several leading sources of the education gap. As this project focuses on the insurance-incentive trade-off and education investments, a natural policy consideration is to promote education investments while maintaining the social safety net. Education-promoting policies aim to improve individuals' productivity and better equip them to be self-sufficient in the labour market. The added benefit from the government's perspective is that more educated and productive individuals will translate into fewer program beneficiaries. In the face of rising fiscal costs of disability policy, addressing such long-run factors to stem flows onto programs is essential in helping sustain the long-run solvency of the disability policy system.

The decomposition exercise provides an understanding of the factors driving the education gap, motivating two types of policy reforms. The first is to reduce the distortionary disincentives related to disability policy. The decomposition exercise found SA-D to be an important contributor to the baseline gap. Reducing this safety net will push people to self-insure their income through investment in their education. The second is to incentivise education by reducing idiosyncratic costs. Given the substantial difference in the costs of post-secondary education for the early-onset group, the government can subsidize the consumption of these individuals during their post-secondary schooling. Such a policy, which takes the form of a scholarship or grant, targets individuals on the margin of choosing post-secondary education who choose s = 0 because of high costs despite a favorable ability draw.

For each scenario, I show the behavioural effects of these policies on individuals' education choices, employment, earnings, consumption, and welfare. I measure welfare implications by calculating the willingness to pay for the new policy through a proportional reduction in consumption at all ages that makes individuals indifferent, ex ante, between the baseline and the policy change considered. It's a consumption equivalent measure interpreted as the percentage of the baseline stream of consumption people are willing to forgo to have the reform in place.⁵⁴ This is obtained by calculating the expected utility at the start of the life-cycle before any uncertainty is resolved. The net impact on the government budget is neutralized in all experiments using a proportional wage tax imposed on all individuals.

7.1 Generosity of SA-D Payments



Figure 5: Proportional Change in SA-D Generosity

The first experiment analyses the effects of scenarios that proportionally change the generosity of SA-D, ranging from 0.8 to 1.2 times its baseline value. Figure 5 plots the simulated changes in individual-level variables (left) and in government revenues and liabilities (right) for early-onset individuals. In each figure, the y-axis depicts the average present value percentage difference in each outcome in each counterfactual

 $[\]overline{{}^{54}\text{The WTP}}$ is calculated as $WTP = (\frac{EV_{baseline}}{EV_{reform}})^{1-\kappa} - 1$. 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.

scenario relative to the baseline, and the x-axis represents the proportional change in SA-D.

First, the incentive cost of SA-D is seen through changes in rates of education and employment. Decreasing the generosity of SA-D to 80% of its baseline value increases the rate of post-secondary attainment by nearly 5%. As SA-D becomes less generous, the relative return to work is greater, which places a higher value on post-secondary education. Employment and earnings increase by nearly 3.5% as individuals choose to work more often on average. These responses result from a direct effect from the relative value of employment to the now less generous SA-D and an indirect effect from more educated individuals receiving higher wage offers.

The effects of increases in SA-D are not symmetric. A 20% increase in SA-D reduces post-secondary schooling by 3.8%, and employment and earnings by approximately 3%. The asymmetry is due to the curvature of utility. The marginal utility of consumption is higher when income is lower, and individuals previously reliant on SA-D are more willing to work despite the utility cost and low earnings. Consumption modestly increases when SA-D is reduced. This occurs because individuals who move into employment under the counterfactual policy tend to earn more than they would receive under the baseline SA-D, reflecting both the utility costs of working and a now relatively higher average education. Lastly, individual welfare is closely tied to SA-D's generosity, as individuals are willing to give up 5% of their ex-ante lifetime consumption in the baseline to maintain SA-D at its baseline level of generosity.

The right side graph considers the impacts of the policy reforms on average present value government liabilities and revenues per early-onset individual. Unsurprisingly, government liability from SA falls as the program becomes less generous. This change combines a direct effect from lower generosity decreasing the relative value of the program and an indirect effect as lower benefits elicit higher rates of post-secondary schooling, raising the relative value of working. The change in DI is positively correlated with the generosity of SA-D. The cost of applying to DI is high, and a more generous SA-D program helps to offset this cost. Hence, this is consistent with these two programs being complements, as in Low and Pistaferri (2015). Tax revenues rise as SA-D decreases, but mildly as the individuals working more in response to lower benefits are typically low earners. In sum, the government saves considerably per early-onset individual relative to the baseline scenario.

In sum, these counterfactuals provide an understanding of the broader incentive costs of SA-D. Lowering SA-D generosity raises rates of post-secondary and promotes employment, lowering costs for the government. However, the welfare of early-onset individuals is closely tied to the generosity of this program, and lowering the generosity of this program comes at a great cost to this group. In addition, the findings suggest that the WTP for individuals without an early-onset disability is positively related to SA-D despite the tax adjustment required to ensure revenue neutrality for the government. Those impacted the most by changes in SA-D lie on the left tail of the income distribution and, therefore, have a larger marginal utility of consumption. This set of experiments highlights the insurance-incentive trade-off of the resources from this program. While behavioural effects are nontrivial, early-onset individuals value the insurance provided considerably.

7.2 Targeted Consumption Subsidy

The second experiment considers a policy that adjusts the consumption of early-onset individuals during post-secondary, for example, via a grant targeting all early-onset individuals. Consumption during post-secondary education in the model is set to the same level for all individuals, and the distribution of

Figure 6: Proportional Change in Consumption During Post-Secondary



idiosyncratic costs calibrates to capture any differences in both pecuniary and non-pecuniary costs across early disability status. As such, these experiments are interpreted in relation to these baseline differences. I consider experiments that proportionally adjust the level of consumption during post-secondary from 0.8 to 1.2 times its baseline level. Figure 6 presents the results of these experiments. Again, the figures plot the average present value percentage difference in each counterfactual scenario relative to the baseline against the proportional change in the policy.

Consider the left figure, which plots the effects on outcomes of early-onset individuals. A subsidy that increases consumption during post-secondary by 20% increases post-secondary attainment by nearly 4%. With more education, early-onset individuals are more productive on average. Earnings rise by 1.8% and push employment to rise by 0.5%. The response curve for post-secondary attainment is concave, which reflects quasi-concavity in preferences. That is, a policy reducing consumption during post-secondary, for instance, by removing existing post-secondary grants for early-onset individuals, elicits a larger change in the rate of post-secondary. The welfare of early-onset individuals increases with the proportional change in consumption during post-secondary.

The right side graph in Figure 6 plots the effects of the policy experiments on the government budget. Increasing individual productivity through incentivising education raises the value of working relative to reliance on disability programs. The most generous policy counterfactual raises education by nearly 4% and reduces government liability for DI and SA-D by 1.3% and 1.8%, respectively. Further, more productive individuals earn more and work more, increasing government average tax revenues per early-onset individual. In sum, a grant increasing consumption in post-secondary by 20% raises government revenues, even when netting out the direct cost of providing the subsidy.

The second class of counterfactual policies increases government expenditures to incentivise post-secondary education, thereby helping to offset the dynamic disincentives embedded in disability policy. In contrast, the first class of policies addresses these disincentives more directly by reducing government spending. Though stylized, the two experiments offer a transparent comparison of the insurance–incentive trade-off inherent in policy reform. Notably, the targeted consumption subsidy mitigates the welfare losses incurred when SA-D generosity is reduced. Furthermore, the behavioural responses induced by these targeted subsidies reduce net costs incurred by the government. However, the response curve for total government revenues is concave, indicating the existence of a threshold beyond which further increases in the grant would lead to declining government revenues.

8 Conclusion

An early-onset disability can impose substantial disadvantages that persist throughout one's life. The effect of an early-onset disability can be mitigated through education investments. However, the incentives to invest in education depend on a number of factors, notably SI policy. This paper develops a structural model to analyze the relative importance of SI policy, among other contributing factors, in affecting observed education investments of individuals with early-onset disabilities. The analysis provides insight into the many ways that an early-onset disability influences education choices and analyzes the role of policy in determining the welfare and outcomes of this population.

The decomposition exercise reveals that disability policy, SA-D in particular, plays a significant role in shaping the gap in educational attainment between individuals with and without an early-onset disability. This finding stems from the fact that the expected value of disability programs reduces the incentive to invest in education by increasing the relative value of the outside option to working. While DI appears to have only a marginal effect on educational decisions, this occurs despite evidence that its expected value is influenced by one's level of schooling. Additionally, the analysis identifies idiosyncratic costs as the leading contributor to the education gap. These costs represent a mix of unobserved psychic and pecuniary factors that influence the net returns to post-secondary education. Unpacking this "black box" to better understand its role in perpetuating educational inequalities among individuals with early-onset disabilities remains an important avenue for future research.

I use the structural model to evaluate the effects of policy reforms on education and labour market outcomes. Two types of reforms are considered: adjustments to the generosity of SA-D benefits and targeted consumption subsidies during post-secondary education. These policy interventions illustrate two approaches to mitigate dynamic disincentives- by removing them or by compensating for them. The findings show that while lower SA-D generosity mitigates the moral hazard in education and employment decisions, it greatly affects individual welfare. Alternatively, targeted consumption subsidies for early-onset individuals can effectively raise post-secondary attainment and while being cost effective for the government. This arises from the fact that those induced to pursue higher education tend to have higher ability and contribute more in tax revenues. The analysis treats the policy environment as exogenously given, and evaluating the optimal design of social insurance while considering broader insurance-incentive trade-offs is left for future research.

References

- Acemoglu, D. and Angrist, J. D. (2001). Consequences of employment protection? the case of the americans with disabilities act. *Journal of Political Economy*, 109(5):915–957.
- Almond, D., Currie, J., and Duque, V. (2018). Childhood circumstances and adult outcomes: Act 11. Journal of Economic Literature, 56(4):1360–1446.
- Altonji, J. G. and Segal, L. M. (1996). Small-sample bias in gmm estimation of covariance structures. Journal of Business and Economic Statistics, 14(3):353–366.
- Ameri, M., Schur, L., Adya, M., Bentley, F. S., McKay, P., and Kruse, D. (2018). The disability employment puzzle: A field experiment on employer hiring behavior. *ILR Review*, 71(2):329–364.
- Attanasio, O. P., Banks, J., Meghir, C., and Weber, G. (1999). Humps and bumps in lifetime consumption. Journal of Business & Economic Statistics, 17(1):22–35.
- Attanasio, O. P. and Weber, G. (1995). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of Political Economy*, 103(6):1121–1157.
- Autor, D. H. (2011). The unsustainable rise of the disability rolls in the united states: Causes, consequences, and policy options. *NBER Working Paper Series-National Bureau of Economic Research*, (17697).
- Autor, D. H. and Duggan, M. G. (2006). The growth in the social security disability rolls: a fiscal crisis unfolding. *Journal of Economic Perspectives*, 20(3):71–96.
- Banks, James, R. B. and Brugiavini, A. (2001). Risk pooling, precautionary saving and consumption growth. *The Review of Economic Studies*, 68(4):757–779.
- Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.
- Bobba, M., Flabbi, L., Levy, S., and Tejada, M. (2021). Labor market search, informality, and on-the-job human capital accumulation. *Journal of Econometrics*, 223(2):433–453.
- Bound, J. and Burkhauser, R. V. (1999). Economic analysis of transfer programs targeted on people with disabilities. *Handbook of Labor Economics*, 3:3417–3528.
- Bound, J., Stinebrickner, T., and Waidmann, T. (2010). Health, economic resources and the work decisions of older men. *Journal of Econometrics*, 156(1):106–129.
- Burkhauser, R. V., Haveman, R. H., and Wolfe, B. L. (1993). How people with disabilities fare when public policies change. Journal of Policy Analysis and Management, 12(2):251–269.
- Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. Journal of Business and Economic Statistics, 21(3):339–353.
- Carneiro, P., Hansen, K. T., and Heckman, J. J. (2003). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college.

- Case, A., Fertig, A., and Paxson, C. (2005). The lasting impact of childhood health and circumstance. Journal of Health Economics, 24(2):365–389.
- Case, A. and Paxson, C. (2010). Causes and consequences of early-life health. Demography, 47(1):S65–S85.
- Cossette, L. (2002). A profile of disability in Canada, 2001, volume 89. Statistics Canada.
- Cunha, F., Heckman, J., and Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, 57(2):191–261.
- Currie, J. (2009). Healthy, wealthy, and wise: Is there a causal relationship between child health and human capital development? *Journal of Economic Literature*, 47(1):87–122.
- Currie, J. and Almond, D. (2011). Human capital development before age five. In Handbook of Labor Economics, volume 4, pages 1315–1486. Elsevier.
- Cutler, D. M., Landrum, M. B., Stewart, K. A., et al. (2006). How do the better educated do it? socioeconomic status and the ability to cope with underlying impairment. *Developments in the Economics of Aging*, pages 203–254.
- Deshpande, M. and Dizon-Ross, R. (2023). The (lack of) anticipatory effects of the social safety net on human capital investment. *American Economic Review*, 113(12):3129–3172.
- Dixon, K. A., Kruse, D., and Van Horn, C. E. (2003). Restricted access: A survey of employers about people with disabilities and lowering barriers to work. John J. Heldrich Center for Workforce Development, Rutgers.
- Druckman, J. N., Levy, J., and Sands, N. (2021). Bias in education disability accommodations. *Economics of Education Review*, 85:102–176.
- Employment and Social Development Canada (2016). Social assistance statistical report: 2009-2013. https://www.canada.ca/en/employment-social-development/services/socialassistance/reports/statistical-2009-2013.html.
- Employment and Social Development Canada Annual report of the canada pen-(2018).fiscal year 2017 to 2018.https://www.canada.ca/en/employment-social- sion plan for development/programs/pensions/reports/annual-2018.html.
- Flinn, C. and Mullins, J. (2015). Labor market search and schooling investment. International Economic Review, 56(2):359–398.
- Gallipoli, G. and Turner, L. (2009). Household responses to individual shocks: Disability and labor supply. https://ageconsearch.umn.edu/record/55323/?v=pdf.
- Golosov, M. and Tsyvinski, A. (2006). Designing optimal disability insurance: A case for asset testing. Journal of political Economy, 114(2):257–279.
- Gourieroux, C., Monfort, A., and Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics*, 8(S1):S85–S118.

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

- Government of Canada (2022). Canada pension plan adjudication framework. https://www.canada.ca/en/services/benefits/publicpensions/cpp/cpp-disability-benefit/medicalprofessionals/framework.html#h2.2.
- Government of Canada (2023). Cpp contribution rates, maximums and exemptions. https://www.canada.ca/en/revenue-agency/services/tax/businesses/topics/payroll/payroll-deductionscontributions/canada-pension-plan-cpp/cpp-contribution-rates-maximums-exemptions.html.
- Grondin, C. (2016). A new survey measure of disability: The Disability Screening Questions (DSQ). Statistics Canada.
- Hartigan, J. A. and Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1):100–108.
- Haveman, R. and Wolfe, B. (2000). The economics of disability and disability policy. Handbook of Health Economics, 1:995–1051.
- Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. Proceedings of the national Academy of Sciences, 104(33):13250–13255.
- Hillel, I. et al. (2020). Holes in the social safety net: Poverty, inequality and social assistance in canada. https://www.csls.ca/reports/csls2020-06.pdf.
- Kellogg, M. D. (2021). Spousal Labor Supply and the Welfare Implications of Disability Insurance Reform. PhD thesis, The University of Chicago.
- Kitao, S. (2014). A life-cycle model of unemployment and disability insurance. Journal of Monetary Economics, 68:1–18.
- Kneebone, R. and White, K. (2015). An overview of social assistance trends in canada. Welfare reform in Canada: Provincial social assistance in comparative perspective, pages 53–92.
- Kostøl, A., Mogstad, M., Setzler, B., et al. (2019). Disability benefits, consumption insurance, and household labor supply. American Economic Review, 109(7):2613–54.
- Liebman, J. B. (2015). Understanding the increase in disability insurance benefit receipt in the united states. Journal of Economic Perspectives, 29(2):123–150.
- Loprest, P. and Maag, E. (2007). The relationship between early disability onset and education and employment. *Journal of Vocational Rehabilitation*, 26(1):49–62.
- Low, H. and Pistaferri, L. (2015). Disability insurance and the dynamics of the incentive insurance trade-off. American Economic Review, 105(10):2986–3029.
- Lundborg, P., Nilsson, A., and Rooth, D.-O. (2014). Adolescent health and adult labor market outcomes. Journal of Health Economics, 37:25–40.

- Maytree-Foundation (2018). Welfare in canada. https://maytree.com/changing-systems/data-measuring/welfare-in-canada.
- Meyer, B. D., Mok, W. K., and Sullivan, J. X. (2009). The under-reporting of transfers in household surveys: its nature and consequences. Technical report, National Bureau of Economic Research.
- Michaud, A. and Wiczer, D. (2018). The disability option: Labor market dynamics with macroeconomic and health risks. Technical report, CHCP Working Paper.
- Milligan, K. (2016). Canadian tax and credit simulator. Database, software and documentation, Version 2016-2.
- Milligan, K. and Schirle, T. (2019). Push and pull: Disability insurance, regional labor markets, and benefit generosity in canada and the united states. *Journal of Labor Economics*, 37(S2):S289–S323.
- Morris, S. P., Fawcett, G., Brisebois, L., and Hughes, J. (2018). A demographic, employment and income profile of canadians with disabilities aged 15 years and over, 2017.
- Navarro, S. and Zhou, J. (2017). Identifying agent's information sets: An application to a lifecycle model of schooling, consumption and labor supply. *Review of Economic Dynamics*, 25:58–92.
- Office of the Auditor General of Canada (2015). Report 6-canada pension plan disability program. https://www.oag-bvg.gc.ca/internet/english/parl_oag_201602_06_e_41063.html.
- Prinz, D., Chernew, M., Cutler, D., and Frakt, A. (2018). Health and economic activity over the lifecycle: Literature review. NBER Working Paper No. 24865.
- Shorrocks, A. F. et al. (2013). Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *Journal of Economic Inequality*, 11(1):99–126.
- Statistics Canada (1993-2022). The canadian tuition and living accommodation costs survey. https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3123.
- Statistics Canada (2012-2018). The longitudinal and international study of adults. all cycles survey and administrative data. https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5144#a1.
- Torjman, S. and Makhoul, A. (2016). *Disability supports and employment policy*. Caledon Institute of Social Policy.

Appendix to "Early-Onset Disability, Education Investments, and Social Insurance"

By Robert Millard

April 4, 2025

1 Model Parameters

Parameter	Description
Individual Heterogeneity	
$ar{a}^{d_0} \ \sigma^2_{a^{d_0}} \ \psi^{d_0} \ \sigma^2_{\psi^{d_0}} \ \sigma^2_{\psi^{d_0}}$	mean of endowed ability distribution variance of endowed ability distribution mean of idiosyncratic cost distribution variance of idiosyncratic cost distribution
Earnings Process	
$\phi^{d_0} \ \mu_1^{d_0,s} \ \mu_2^{d_0,s} \ \sigma_{\xi_0}^2 \ \sigma_{\xi_0}^2 \ h_{d_0}^s$	direct effect of disability on earnings return to potential experience return to potential experience squared variance of productivity shock distribution mean of initial productivity shock distribution variance of initial productivity shock distribution return to post-secondary
Utility Parameters	
$\beta \\ \kappa \\ \theta \\ \eta_1 \\ \eta_2 \\ F^d$	discount factor coefficient of relative risk aversion utility cost of disability utility cost of working utility cost of working with a disability cost of working with a disability at old age
Policy Parameters	
$\pi^s \ \pi^{SA} \ C^{s,d_0}_{App}$	probability of DI acceptance probability of SA-D acceptance utility cost of DI application
Labour Market Environment	
$\delta^{d,s}_t \\ \lambda^{s,d}_t \\ \gamma^{d_0,t}_{i,j}$	exogenous job destruction rate exogenous job arrival rate disability transition probability

Table 1: Summary of Model Parameters.

2 Measuring Disability in the Data

Disability in the model is measured by reported limitations to activities of daily living (LADL). The set of LADLs is derived from a short version of a module called "the disability screening questions" developed by Statistics Canada for identifying individuals with disabilities in general population surveys (Grondin, 2016). This model distinguishes five main areas of activity limitation: Seeing, Hearing, Physical, Cognitive, and Mental Health. Sample survey questions used to identify disability status in the data are reported in Table 2

Table 2: Survey Questions on Limitations to Daily Activities

Physical limitation

-How much difficulty do you have walking on a flat surface for 15 minutes without resting?

-How much difficulty do you have walking up or down a flight of stairs, about 12 steps without resting?

-How much difficulty do you have reaching in any direction, for example, above your head?

-How much difficulty do you have using your fingers to grasp small objects like a pencil or scissors?

-Do you have pain that is always present?

Cognitive limitation

-Do you think you have a condition that makes it difficult in general for you to learn? This may include learning disabilities such as dyslexia, hyperactivity, attention problems, etc..

-Has a teacher, doctor or other health care professional ever said that you had a learning disability?

-Has a doctor, psychologist or other health care professional ever said that you had a developmental disability or disorder? This may include Down syndrome, autism, Asperger syndrome, mental impairment due to lack of oxygen at birth, etc..

-Do you have any ongoing memory problems or periods of confusion? Please exclude occasional forgetfulness such as not remembering where you put your keys.

Mental Health limitation

-Do you have any emotional, psychological or mental health conditions? These may include anxiety, depression, bipolar disorder, substance abuse, anorexia, etc..

Sensory Limitation

- How often does this difficulty seeing limit your daily activities?
- How often does this difficulty hearing limit your daily activities?

Note: Provides the disability-related survey questions in LISA used to construct disability status.

2.1 Validity of Disability Measures

Much research in health economics has focused on the validity of self-reported measures of one's health. One concern relates to the inherent subjectivity of how one assesses one's own health. For example, two otherwise identical individuals may differ in the reported severity of their disability. Additionally, critics of self-reported health measures argue that individuals may exaggerate the existence or severity of their health condition to justify poor economic outcomes or attachment to government programs, a phenomenon referred to as justification bias. The evidence on the endogeneity of self-reported health measures and the extent of measurement error are mixed (Black et al., 2017). Although, it is important to note that recent articles tend to find evidence for state-dependent reporting.¹

My disability measure is derived from a respondent reporting any positive limitations to a specified activity and abstracts from the degree of impairment. This approach mitigates concerns related to subjectivity in the scale of impairment from a self-reported activity limitation, as I do not distinguish conditions along the severity margin. Moreover, much of the evidence on justification bias is based on broad questions about one's health or disability, such as "do you have a medical or physiological condition that impairs the type or amount of work you can do." The questions about activity limitations in this survey are linked to specific tasks, such as walking on a flat surface for fifteen minutes, grasping a small object like scissors, or experiencing ongoing memory problems or periods of confusion. Additionally, the presence of some activity limitations is elicited based on whether the respondent has been diagnosed with a specific condition, such as a learning or developmental disorder, by a healthcare professional.² Last, mental health is identified using specific examples of diagnoses, such as anxiety, depression, bipolar disorder, or anorexia. These approaches narrow the scope of justification bias to be anchored to the activities in question, base the existence of a limiting condition on the diagnosis of a medical professional, or frame limitations related to mental health with specific examples of diagnoses. I follow much of the related literature and take the responses to questions on limitations to daily activities as given. However, I acknowledge the empirical concerns that are inherent to any self-reported measures of health.

¹It has been found that self-reported disability is close to exogenous, may actually under-represent the extent disabled population, and may even underestimate the true impact of disability on relevant labour market outcomes (Stern, 1989; Bound and Burkhauser, 1999; Burkhauser et al., 2002). Others have found evidence of justification bias related to labour market states inflating the prevalence of health conditions (Benítez-Silva et al., 2004; Baker et al., 2004; Black et al., 2017). Moreover, alternate approaches to identify individuals with disabilities, for instance, by using disability insurance beneficiaries to define the population with a disability, have been found to under-represent the population of individuals who are limited enough in the labour market to be classified as "disabled" (Bound, 1989)

 $^{^{2}}$ This type of question has been used to assess the validity of self-reported health measures in Baker et al. (2004)

3 Value Functions and Numerical Solution to Structural Model

There is no analytical solution to the model so it is solved numerically. For a given set of the structural parameters, 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 in T (retirement at age 65), I iterate backward, numerically approximating the value functions, characterizing the work decision and Disability Insurance (DI) application decision at each age after eighteen as a function of $S_t = \{d_t, \epsilon_t, e_{t-1}, \rho_t\}$. Given the solution to the individual's labour market decisions, I solve the policy function for the education choice at age eighteen as a function of initial heterogeneity, $\{a, d_0, \psi\}$.

Retirement

Solving the model starts with the terminal condition, retirement. The value of the terminal period is deterministic for a given set of the state variables. I assume that state variables remain fixed as soon as an individual retires, $S_t = S_{t+1} = \overline{S} = \{\overline{d}, \overline{\epsilon}, \overline{e}, \overline{\rho}\}$. 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.³ I assume individuals expect retirement to last until age 75, after which they die with certainty. The value of retirement is

$$V_t^R(\bar{S}) = u^N(c_t; \bar{d}) + \beta V_{t+1}^R(\bar{S})$$
(1)

$$= u^{N}(\bar{c}; \bar{d}) + \sum_{\tau=1}^{T^{L}} \beta^{\tau} u^{N}(\bar{c}; \bar{d})$$
(2)

s.t.
$$c_t = 5500 + 0.25\bar{e}.$$
 (3)

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 do not have the option to retire.

Value of Working

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

$$V_t^E(S_t) = u^W(c_t; d_t) + \beta E_t \left[\delta_t^{d_0, s} V_{t+1}^U(S_{t+1}) + (1 - \delta_t^{d_0, s}) \max \left\{ V_{t+1}^U(S_{t+1}), V_{t+1}^E(S_{t+1}) \right\} \right]$$
(4)

s.t.
$$c_t = \tau (W_t L_t, 0) + S A_t (\tau (W_t L_t, 0), d_t),$$
 (5)

$$e_t = f(e_{t-1}, W_t, t).$$
 (6)

³The individual's contribution period ends at T^L so their earnings index remains constant after this time.

Value of Not Working and Receiving Social Assistance (SA)

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

$$V_t^U(S_t) = u^N(c_t; d_t) + \beta \mathop{\rm E_tmax}_{m_t \in \{0,1\}} \left[m_t \left(\pi^s V_{t+1}^{DI}(S_{t+1}) + (1 - \pi^s) V_{t+1}^U(S_{t+1}) - C_{app}^{d_0,s} \right) \right]$$
(7)

$$+ (1 - m_t) \Big(\lambda_t^{d_0, s} \max \Big\{ V_{t+1}^U(S_{t+1}), V_{t+1}^E(S_{t+1}) \Big\} + (1 - \lambda_t^{d_0, s}) V_{t+1}^U(S_{t+1}) \Big) \Big]$$
(8)

s.t.
$$c_t = SA(0, d_t),$$
 (9)

$$e_t = f(e_{t-1}, 0, t). (10)$$

DI Beneficiary

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

$$V_t^{DI}(S_t) = u^N(c_t; d_t) + \beta E_t \bigg[(1 - \lambda_t^{d_0, s}) \max\{V^U(S_{t+1}), V^{DI}(S_{t+1})\}$$
(11)

$$+ \lambda_t^{d_0,s} \max\{V^E(S_{t+1}), V^U(S_{t+1}), V^{DI}(S_{t+1})\}$$
(12)

s.t.
$$c_t = \tau(0, DI_t) + SA_t(\tau(0, DI_t,), d_t)$$
 (13)

$$e_t = e_{t-1}.\tag{14}$$

In each period t, for every possible combination of the discrete state variables—both those that are time-varying and those that are fixed—I evaluate the continuation value (Emax) on a discretized grid of the continuous state variables. The continuous state variables are initially $(a_i, \epsilon_{it}, e_{i,t-1})$, where a_i is endowed ability, ϵ_{it} is the accumulated productivity shock, and $e_{i,t-1}$ is the earnings index. Because a_i and ϵ_{it} only affect future earnings, I can reduce the problem to tracking $(W_{it}, e_{i,t-1})$ where W_{it} is the current earnings.

To compute the expected continuation value, I integrate out the next period's productivity shock, ξ_{it+1}^{s,d_0} . Assuming this shock is normally distributed, I use Gauss-Hermite quadrature to numerically approximate the integral over its distribution. For each realization of the discrete state variables, I construct an approximation of the continuation value by evaluating the expected payoff on a discrete grid for $(W_{it}, e_{i,t-1})$. Finally, to handle points that lie between the grid values in the continuous space, I apply bilinear interpolation. This approach ensures a smooth approximation of the continuation value function while keeping the computational burden tractable.

3.1 Smoothing

Applying indirect inference to a discrete choice model presents challenges due to the discontinuous nature of the mapping from structural parameters to simulated data. Small changes in the structural parameters can lead to abrupt shifts in the simulated outcomes, causing the auxiliary model's parameter estimates to change discontinuously. These discrete jumps introduce discontinuities in the objective function, complicating optimization. Additionally, some parameter changes may not affect the discrete choices at all, resulting in flat regions in the objective function.

To address these issues, I adopt a Generalized Indirect Inference (GII) procedure, which smooths the objective function and mitigates both flat spots and discontinuities (Bruins et al., 2018), (Keane and Smith, 2003). The key idea is to apply distinct auxiliary models to the simulated and observed data. In particular, the auxiliary model for the simulated data is designed to fit the continuous latent variables that underlie the observed discrete outcomes. Provided that both auxiliary models yield asymptotically equivalent vectors of pseudo-true parameters, the GII estimator—defined by minimizing the distance between the two models—remains consistent and asymptotically normal.

To implement GII, I introduce an i.i.d. taste shock, $\zeta_t^k = (\zeta_t^E, \zeta_t^U, \zeta_t^{DI})$, into the utility associated with each labor market state. These shocks are interpreted structurally as unobserved state variables known to the agents but not to the econometrician. The shocks follow a multivariate extreme value distribution with scale parameter λ . Their inclusion necessitates modifications to both the model's solution method and the estimation algorithm.

In solving the model, I follow a similar procedure as previously described, with the key distinction that I now account for the newly introduced state variables when computing the expected maximum (Emax) functions within the continuation values at each decision point. To illustrate, the value function for the employed state becomes:

$$V_t^E(S_t) = u^W(c_t; d_t) + \lambda \zeta_t^E + \beta E_t \bigg[\delta_t^{d_0, s}(V_{t+1}^U(S_{t+1}) + \lambda \zeta_{t+1}^U)$$
(15)

$$+ (1 - \delta_t^{d_0,s}) \max \left\{ V_{t+1}^U(S_{t+1}) + \lambda \zeta_{t+1}^U, V_{t+1}^E(S_{t+1}) + \lambda \zeta_{t+1}^E \right\}$$
(16)

$$V_t^E(S_t) = u^W(c_t; d_t) + \lambda \zeta_t^E + \beta \left[\delta_t^{d_0, s} E_t(V_{t+1}^U(S_{t+1}) + \lambda \zeta_{t+1}^U) \right]$$
(17)

$$+ (1 - \delta_t^{d_0,s}) \mathcal{LS}_{k \in E, U} (V_{t+1}^U(S_{t+1}), V_{t+1}^E(S_{t+1}))$$
(18)

where \mathcal{LS} is the log-sum function

$$\mathcal{LS}_{k \in E, U} \left(V_{t+1}^U(S_{t+1}), V_{t+1}^E(S_{t+1}) \right) = \lambda \log \left(\exp(V_{t+1}^U(S_{t+1})/\lambda) + \exp(V_{t+1}^E(S_{t+1})/\lambda) \right).$$
(19)

Now, the conditional choice probability of each labor market state at period t is given by,

$$Pr(V = V_t^j | S_t) = \frac{exp(V_t^j / \lambda)}{exp(V_t^U / \lambda) + exp(V_t^E / \lambda)}$$
(20)

The estimation procedure for the model with taste shocks follows similar steps as before, with one key modification: moments in the auxiliary model are now calculated using choice probabilities rather than observed outcomes. For example, the auxiliary model includes conditional employment rates computed from the observed data. I estimate the model's parameters by matching these observed rates to the corresponding conditional employment probabilities generated by the simulated model. The fundamental principle of Generalized Indirect Inference (GII) is that the estimation procedures applied to the observed and simulated data need not be identical, so long as both yield consistent estimates of the same vector of pseudo-true parameter values.

4 Censoring

The data used in the analysis is an unbalanced panel, as such there is considerable censorship present when calculating the moments making up the auxiliary model for estimation. To address this, I replicate censoring observed in the data and impose it when calculating the moments using data simulated from the model. I calculate the probability of an observation being censored conditional on age, a lag for censorship in the previous period (L.1), censorship in the previous two periods (L.2), and censorship in the previous three periods (L.3). I estimate a separate linear probability model conditional on d_0 and s, giving four sets of estimates. The results from the estimation are reported in Table 3 below.

	Not Early		Early	
	Low Educ	\mathbf{PS}	Low Educ	\mathbf{PS}
age	-0.001	0	0	-0.001
	(0)	(0)	(0)	(0)
L.1	0.358	0.384	0.333	0.302
	(0.021)	(0.014)	(0.065)	(0.05)
L.2	0.154	0.137	0.122	0.183
	(0.041)	(0.027)	(0.127)	(0.103)
L.3	0.142	0.159	0.228	0.214
	(0.043)	(0.028)	(0.138)	(0.108)
Intercept	0.059	0.039	0.046	0.079
-	(0.005)	(0.003)	(0.016)	(0.012)

Table 3: Censoring

Note: Reports point estimates used to construct probabilities of censoring in the panel data across disability and education subgroups, accounting for lag structures. Standard errors of estimates reported in brackets below point estimates.

5 Descriptive Statistics

Early-Onset	0.460
•	(0.037)
Not Early Disabled	0.640

Table 4: Likelihood of Post-Secondary Attainment by Early Disability Status

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

Table 4 shows that the likelihood of completing post-secondary is 18 percentage-points lower for earlyonset individuals. Less than half of individuals affected by an early-onset disability complete a post-secondary degree.

	Not Early	y Disabled	Early-Onset		
A1137 • TI 36 1	Low Education	Post-Secondary	Low Education	Post-Secondary	
All Years in Labour Market					
Annual Earnings(\$)	32300	50900	26000	40400	
	(21300)	(31600)	(19900)	(27400)	
Employment Rate	0.740	0.846	0.508	0.753	
First 3 years in Labour Market					
Annual Earnings (\$)	15100	20700	12900	18200	
· · · · · · · · · · · · · · · · · · ·	(10800)	(14300)	(10000)	(13300)	
Employment Rate	0.810	0.862	0.579	0.815	

Table 5: Employment and Earnings by Education Level and Early Disability Status.

Notes: Estimates are from T1FF years 1989-2016 and survey weights applied to represent the of Canada population in 2012. Standard deviations are reported in parenthesis below.

Table 5 presents statistics on lifetime earnings and employment by early disability status and education level. Individuals with early-onset disabilities and low education who are employed earn approximately 20% less than their counterparts without early-onset disabilities, increasing their risk of applying for Social Insurance (SI). These lower returns to work are reflected in significantly reduced lifetime employment rates—23 percentage points lower than those of similarly educated individuals without early disabilities. The third and fourth rows of Table 5 report average earnings and employment in the first three years following labor market entry. The observed differences by early disability status within the low education group are smaller in this early period, suggesting that early-onset disabilities may hinder the accumulation of skills over time. Among those with post-secondary education, the earnings gap by early disability status is comparable in magnitude to that observed in the low education group. Early-onset individuals with post-secondary education earn about 20% less than their non-disabled peers, potentially reflecting lower average ability, reduced financial returns to education, or both. However, their average employment rates are much closer to those of non-disabled individuals, indicating relatively higher returns to work within this subgroup.

	Not Early	Disabled	Early	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)	
$Age \ge 45$	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)	
$Age \ge 45$	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)	

Table 6: Average Rate and Transfer Amount From Social Assistance (SA) and Disability Insurance (DI) by Education Level and Early Disability Status

Notes: Estimates are from T1FF years 1989-2016 and survey weights applied to represent the of Canada population in 2012. Standard deviations are reported in parenthesis below.

Table 6 presents statistics on the likelihood of receiving transfers and the average benefit amounts from Disability Insurance (DI) and Social Assistance (SA), disaggregated by early disability status and education level. The first two rows indicate that individuals with early-onset disabilities are substantially more likely to receive SA benefits early in life and, on average, receive larger transfers. Across all education levels, the proportion of individuals who ever become SA recipients is more than double for the early-onset group. Notably, over 30% of early-onset individuals with low education depend on SA at some point during their lives. Rows 3 and 4 show that the difference in average SA benefits received between early-onset and non-disabled individuals decreases with education. Early-onset individuals with low education receive approximately \$2,000 more per year in SA benefits, compared to a difference of around \$800 for those with post-secondary education

Rows 5 and 6 report that the likelihood of receiving Canada Pension Plan Disability (CPP-D) benefits is relatively low, with approximately 4% of early-onset individuals eventually becoming beneficiaries. It is important to interpret this figure as representing only those who both applied for and were accepted into the CPP-D program. In practice, many more individuals may apply but are denied; for instance, in 2014–2015, only 43% of CPP-D applications were approved (Office of the Auditor General of Canada, 2015). Lastly, the average size of DI benefits increases with age, reflecting growth in lifetime earnings.

6 Model Policy Environment

6.1 Tax Environment

Parameters for the income tax brackets and marginal tax rates were derived from the Canadian Tax and Transfer Simulator (Milligan, 2016). In each province and calendar year, I cap the upper threshold to tax brackets to give me 5 distinct tax brackets. I then calculate the economy's average income brackets and marginal tax rates across all years and provinces in the support of my data. I each province-year tax regime based on the joint density of calendar year and province in my sample. Table 7 reports the resulting tax system used in the model.

Table 7:	Tax	Brackets	and	Marginal	Tax	Rates.
T 00010 11		1010011000	COLL OF	THE GILLON		1000000

Income Bracket	Tax Rate
	$\begin{array}{c} 0.2280 \\ 0.2944 \\ 0.3433 \\ 0.3621 \\ 0.3833 \end{array}$
L / J	

Shows the tax schedule implemented in the model based on bracketed income thresholds and associated rates. Tax rates and income brackets derived using parameters of the Canadian Tax and Transfer Simulator (Milligan, 2016).

6.2 Social Assistance Regimes

In Canada, SA policies vary across provinces and over calendar time. For each province and each time period, I represent the SA policy as a two-element "couplet" showing the maximum benefit available under SA and under SA-D. Given 10 provinces observed over 29 periods, this yields 290 distinct couplets. Accommodating three hundred different SA policies is computationally intractable. To simplify, I group "similar" couplets using a k-means clustering algorithm (Hartigan–Wong), which clusters observations based

on Euclidean distances (Hartigan and Wong, 1979). The algorithm partitions the 290 couplets into clusters by minimizing the sum of squared distances between points and their assigned cluster centers.

Hartigan–Wong algorithm proceeds by trying to place each data point into the "best" cluster, which loosely translates to minimizing the overall within-cluster variance (the total euclidean distance of points to their cluster centers). Given a set of data points (290 two-dimensional "couplets" in my application). I also decide on a number of clusters k=2. The algorithm begins by assigning each point to one of k clusters in some initial way (often randomly). After the initial grouping, the algorithm checks whether moving each point it from its current cluster to a different cluster would reduce the overall distance within all clusters. If it finds that moving a point to a different cluster yields a lower overall sum of squared distances, it makes that move. Each time a point is reassigned, the center (mean) of both the old cluster and the new cluster is updated to reflect the change. The algorithm continues through the points and reassigning them whenever a beneficial move is found. Once no further improvements can be found (in terms of reducing overall distance), the algorithm has converged.



Figure 1: SA Regimes and Clusters by Province and Year

Note: Graph illustrates the k-means clustering of social assistance policies across province-time pairs. The generosity of regular SA-D is reported on the horizontal axis and generosity of SA on the vertical axis. Each point represents a province-year SA regime. Regimes are grouped into low generosity (circles) and high generosity (triangles) regimes.

Figure 1 illustrates these clusters: each point corresponds to a province-time couplet, shaded regions show which couplets are grouped together, and each black dot marks the cluster center (a weighted average of its members). The cluster centers are the SA regimes used in the model. I choose two clusters: one representing less generous SA policies and another representing more generous ones.

7 Estimation of search frictions

To calculate job arrival rates, I first estimate parameters estimates from the following probit model:

$$UE_{i} = \gamma_{0}^{s,d_{0}} + \gamma_{1}^{s,d_{0}}LS_{i} + \gamma_{2}^{s,d_{0}}age_{i} + LS_{i} * age_{i}\gamma_{3}^{s,d_{0}} + \epsilon_{i},$$
(21)

where the dependent variable, UE_i is an indicator equal to one if an *i* is employed (part-time or full-time). The variable LS_i indicates whether the individual was actively searching for a job in the previous month. Probit regressions are estimated separately by schooling level (s) and early-onset disability status (d_0). Using the estimated coefficients, I calculate the marginal effect of job search on employment probability across age and convert monthly arrival rates into annual equivalents.

To calculate job destruction rate, I first obtain parameter estimates from the following model,

$$EU_i = \beta_0^{s,d0} + \beta_1^{s_d0} age_i + \epsilon_i. \tag{22}$$

where the dependent variable, EU_i , equals one if the individual was fired or laid off since the last survey wave. As before, the model is estimated separately by s and d_0 , and I use the resulting estimates to predict age-specific separation probabilities.

	d0s0	d0s1	d1s0	d1s1
LS	-0.728	-1.001	-0.619	-0.578
	(0.104)	(0.107)	(0.233)	(0.301)
age	-0.001	-0.006	-0.006	-0.011
	(0)	(0)	(0.001)	(0.001)
LS^*age	-0.014	-0.012	-0.017	-0.02
	(0.003)	(0.002)	(0.009)	(0.008)
Intercept	0.082	0.507	-0.128	0.527
	(0.01)	(0.01)	(0.026)	(0.033)
	. ,	. ,	. ,	,
age	-0.006	-0.011	-0.031	-0.016
	(0.003)	(0.003)	(0.008)	(0.009)
Intercept	-0.958	-0.984	-0.129	-0.514
	(0.137)	(0.12)	(0.277)	(0.358)

Table 8: Models for Job Arrival Rate and Destruction Rate

Note: Table provides estimated coefficients from probit models of job arrival and separation, by education level and disability status. Standard errors are reported in brackets below point estimates.

Table 8 presents the estimates from estimating (17) and (18). The estimation results indicate that individuals with post-secondary education (s = 1) receive job offers at a higher rate. Conditional on schooling, individuals with early-onset disabilities are less likely to receive job offers—consistent with employer perceptions of lower productivity, higher accommodation costs, or bias against hiring individuals with disabilities (Dixon et al., 2003). In contrast, job separation rates are lower for individuals with higher education and higher for those with early-onset disabilities, conditional on education. This aligns with the interpretation that more stable, permanent jobs are available to individuals with post-secondary credentials.

8 Description of Decomposition

I follow the framework of Shorrocks et al. (2013) in decomposing the education gap into its contributing components. Consider a statistical indicator, I, which can be fully expressed as a function of m contributory factors,

$$I = f(X_1, ..., X_m). (23)$$

In this application, I is the education differential between individuals with and without an early onset disability. The contributing factors, X_m , are sets of structural parameters that differ by initial disability status, and the function f is the mapping from the structural model to the education differential. Let F(S)be the value of I when a set of factors $X_k, k \notin S$, have been shut off. A decomposition of the model structure $\{K, F\}$ is defined as a set of real values $C_k, k \in K$, representing the contribution of each factor. That is, the contribution of a factor corresponds to the change in I when that factor was shut off. A decomposition rule is a function that generates these factor contributions:

$$C_k = C_k(K, F) \tag{24}$$

The first decomposition I implement calculates the marginal impact on the education gap when shutting down a single factor, with all other factors on. This is given by:

$$C_k(K,F) = F(K) - F(K/\{k\}), k \in K$$
(25)

This decomposition represents the *ceteris paribus* effect of each contributing factor on the gap, holding all other model features constant. However, the individual contributions derived from this method do not, in general, sum to reproduce the entire baseline education gap.

As an alternative, I employ the Shapley decomposition, which ensures that the sum of the factor contributions equals the total baseline gap. The Shapley decomposition is calculated based on calculating the marginal impact of each factor across all m! possible ordered sequences in which the factors could be eliminated:

$$C_j = \sum_{k=0}^{n-1} \frac{(n-k-1)!k!}{n!} \left(\sum_{s \in S_k / \{X_j\}: |s|=k} \left[f(s \cup X_j) - f(s) \right] \right)$$
(26)

where n is the total number of arguments in the original function, and $S_k/\{X_j\}$ is the set of all "submodels" of size k that exclude factor X_j . The weighting term, $\frac{(n-k-1)!k!}{n!}$ reflects the probability that a particular submodel of size k is randomly selected under uniform permutation.

The Shapley decomposition has three desirable properties. First, it is exact—the contributions of all factors sum to match the total education gap. Second, it satisfies symmetry: if two factors have identical marginal effects across all permutations, their contributions will be equal. This property ensures path independence in the estimation of each factor's contribution.⁴ Third, the method accommodates hierarchical structures, allowing for decomposition into both primary and secondary contributing factors.

⁴In contrast, sequential shutdown methods can be path-dependent.

9 Auxiliary Model and Fit of Moments

Tables 9 - 19 display the full set of auxiliary moments used in estimation. Each table reports the moments calculated in the observed data, the moments calculated using data simulated with the model, and the standard error of the observed data moment. Estimation consists of 217 moments, including education rates and regressions (Table 9), coefficients from DI rate and DI flow regressions (Tables 10, 11, and 12), employment rates, flows, and regressions (Tables 13 and 14), and earnings distributions and regressions (Tables 15, 16, 17, 18, and and 19).

Moment	Data	Simulation	Standard Error	•
	Educe	ation Pata		
	Educa	ation nate		
$Frac(s = 1 d_0 = 1)$	0.466	0.457	0.037	
$Frac(s = 1 d_0 = 0)$	0.658	0.638	0.012	
(I * /				
Li	near Pro	bability Mode	els	
d_0	-0.125	-0.094	0.033	
Ŷ	0.146	0.254	0.015	
Intercept	-0.673	-1.736	0.143	
σ_ψ^z	0.214	0.220	0.003	
d	0.804	1.059	0.294	
\hat{u}_0	-0.694	1.056	0.304	
V m X û	0.130	0.330 0.125	0.010	
$pr \times v$	0.000	-0.125	0.042 0.154	
$_{-2}$	-0.084	-2.499	0.104	
O_{ψ}	0.215	0.219	0.005	
Condition	nal Linea	ar Probability	Models	
	0			
Conditional on d_0	0 = 0	0.015	0.007	
\hat{pr}	-0.886	-0.215	0.307	
Ŷ	0.098	0.326	0.022	
pr imes v	0.094	0.014	0.033	
Intercept	-0.230	-2.359	0.201	
σ_ψ^2	0.211	0.206	0.004	
Conditional on d_0	= 1			
pr	-1.131	-0.183	0.737	
ŷ	0.169	0.205	0.049	
$pr imes \hat{v}$	0.121	0.009	0.081	
Intercept	-0.989	-1.330	0.450	
σ_{sb}^2	0.224	0.228	0.009	
Ψ				

Table 9: Education Rates and Regressions

Moment	Data	Simulation	Standard Error
	d_0	= 0, s = 0	
age	0.007	0.011	0.002
age^2	0.000	0.000	0.000
age^3	0.000	0.000	0.000
Intercept	-0.063	-0.105	0.027
	d_0	0 = 0, s = 1	
age	0.006	0.005	0.002
age^2	0.000	0.000	0.000
age^3	0.000	0.000	0.000
Intercept	-0.068	-0.053	0.018
	d_0	s = 1, s = 0	
age	0.021	0.016	0.014
age^2	-0.001	-0.001	0.000
age^3	0.000	0.000	0.000
Intercept	-0.193	-0.141	0.153
	d_0	s = 1, s = 1	
age	-0.003	0.006	0.014
age^2	0.000	0.000	0.000
age^3	0.000	0.000	0.000
Intercept	0.038	-0.045	0.167

Table 10: OLS Regression Coefficients: DI Rate

=

Moment	Data	Simulation	Standard Error
	d_0	0 = 0, s = 0	
age	-0.002	0.000	0.001
age^2	0.000	0.000	0.000
age^3	0.000	0.000	0.000
Intercept	0.027	0.007	0.010
	de	-0 e - 1	
	0.001	$\frac{0}{0} = 0, 3 = 1$	0.001
uge	0.001	0.001	0.001
age^3	0.000	0.000	0.000
Intercept	-0.007	-0.013	0.008
		0.010	
	$\underline{d_0}$	s = 1, s = 0	
age	-0.004	-0.008	0.004
age^2	0.000	0.000	0.000
age^3	0.000	0.000	0.000
Intercept	0.045	0.106	0.042
	d_{0}	s = 1, s = 1	
aae	0.005	-0.006	0.006
age^2	0.000	0.000	0.000
age^3	0.000	0.000	0.000
Intercept	-0.058	0.069	0.070
· · · · · · · · · · · · · · · · · · ·			

Table 11: OLS Regression Coefficients: DI Flow

=

Table 12: OLS Regression Coefficients: Pre-DI Ln(Average Earnings) and Employment

Moment	Data	Simulation	Standard Error
Depen	dent Var	iable : $Ln(Av$	erage Earnings)
d_0	-0.119	-0.216	0.300
s	0.209	-0.038	0.106
$s \times d_0$	-0.316	0.348	0.390
intercept	10.141	9.795	0.074
Deper	ndent Va	riable: Averag	ge Employment
d_0	-0.280	-0.407	0.173
s	0.104	0.169	0.059
$s \times d_0$	0.034	0.153	0.213
intercept	0.706	0.583	0.044

Notes: dependent Variables are calculated as the average over the 5 periods prior to applying for DI.

Moment	Data	Simulation	Standard Error
- ·			
Employme	nt Rates	3	0.000
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 0, s_i = 0, t < 45)$	0.874	0.785	0.003
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 0, s_i = 0, t \ge 45)$	0.797	0.773	0.004
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 0, s_i = 1, t < 45)$	0.908	0.879	0.002
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 0, s_i = 1, t \ge 45)$	0.850	0.892	0.003
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 1, s_i = 0, t < 45)$	0.670	0.819	0.015
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 1, s_i = 0, t \ge 45)$	0.479	0.678	0.009
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 1, s_i = 1, t < 45)$	0.831	0.892	0.008
$Fr(L_{it} = 1 d_0 = 0, d_{it}^* = 1, s_i = 1, t \ge 45)$	0.638	0.863	0.007
$Fr(I_{11} - 1 d_1 - 1 d_2 - 1 c_1 - 1 c_2 - 0 t < 45)$	0 591	0 507	0.014
$Fr(L_{it} - 1 d_0 - 1, u_{it} - 1, s_i - 0, t < 45)$ $Fr(L_{it} - 1 d_0 - 1, d^* - 1, s_i - 0, t > 45)$	0.521	0.307	0.014
$Fr(L_{it} - 1 d_0 - 1, u_{it} - 1, s_i - 0, t \ge 45)$ $Fr(L_{it} - 1 d_0 - 1, d^* - 1, s_i - 1, t < 45)$	0.400	0.425 0.754	0.020
$Fr(L_{it} - 1 d_0 - 1, d_{it} - 1, s_i - 1, t < 45)$ $Fr(L_{it} - 1 d_0 - 1, d^* - 1, s_i - 1, t < 45)$	0.611	0.754	0.009
$FT(L_{it} = 1 a_0 = 1, a_{it} = 1, s_i = 1, t \ge 45)$	0.011	0.055	0.018
Employment Tra	ansition	Rates	
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 0, \overline{s_i = 0, t < 45})$	0.046	0.123	0.002
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 0, s_i = 0, t \ge 45)$	0.040	0.089	0.002
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 0, s_i = 1, t < 45)$	0.037	0.076	0.001
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 0, s_i = 1, t \ge 45)$	0.039	0.060	0.001
$E_{\rm m}(I = 1 I = 0 d = 0 c = 0 d < 45)$	0.050	0 129	0.009
$Fr(L_{it} = 1 L_{it} = 0, d_{i0} = 0, s_i = 0, t < 45)$	0.050	0.152	0.002
$Fr(L_{it} = 1 L_{it} = 0, d_{i0} = 0, s_i = 0, t \ge 45)$	0.028	0.074	0.002
$Fr(L_{it} = 1 L_{it} = 0, d_{i0} = 0, s_i = 1, t < 45)$	0.047	0.115	0.001
$FT(L_{it} = 1 L_{it} = 0, a_{i0} = 0, s_i = 1, t \ge 45)$	0.024	0.055	0.001
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 1, s_i = 0, t < 45)$	0.080	0.139	0.007
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 1, s_i = 0, t \ge 45)$	0.029	0.049	0.006
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 1, s_i = 1, t < 45)$	0.050	0.127	0.005
$Fr(L_{it} = 0 L_{it} = 1, d_{i0} = 1, s_i = 1, t \ge 45)$	0.045	0.081	0.008
$Fr(L_{11} - 1 L_{11} - 0, d_{12} - 1, s_1 - 0, t < 45)$	0.067	0.150	0.007
$Fr(L_{it} - 1 L_{it} = 0, a_{i0} = 1, s_i = 0, t < 45)$ $Fr(L_{it} - 1 L_{it} = 0, d_{i0} = 1, s_i = 0, t > 45)$	0.007	0.130	0.007
$Fr(L_{it} = 1 L_{it} = 0, a_{i0} = 1, s_i = 0, t \ge 45)$ $Fr(L_{it} = 1 L_{it} = 0, d_{it} = 1, s_i = 1, t \le 45)$	0.019	0.055	0.005
$Fr(L_{it} = 1 L_{it} = 0, a_{i0} = 1, s_i = 1, t < 45)$ $Fr(L_{it} = 1 L_{it} = 0, d_{it} = 1, s_i = 1, t < 45)$	0.038 0.027	0.102	0.005
$FT(L_{it} = 1 L_{it} = 0, a_{i0} = 1, s_i = 1, t \ge 43)$	0.027	0.001	0.005
Employment Rate at L	abour N	Iarket Entry	
$Fr(L_{it} = 1 d_{i0} = 0, \overline{s_i = 0, t = 1})$	0.809	0.648	0.009
$Fr(L_{it} = 1 d_{i0} = 0, s_i = 1, t = 4)$	0.862	0.822	0.006
$Fr(L_{it} = 1 d_{i0} = 1, s_i = 0, t = 1)$	0.579	0.363	0.028
$Fr(L_{it} = 1 d_{i0} = 1, s_i = 1, t = 4)$	0.815	0.552	0.022

Table 13: Conditional Employment Rates and Flows

$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	Table 14:	OLS Reg	gression Coeffi	cients: Employment
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	Moment	Data	Simulation	Standard Error
$\begin{array}{cccccccccccccccccccccccccccccccccccc$				
$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$			$\underline{d_0 = 0, s = 0}$	
$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	age	-0.034	0.040	0.007
$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	$age^{2}/100$	0.001	0.000	0.000
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$age^{3}/100$	0.000	0.000	0.000
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	d^*	-0.263	-0.019	0.009
$\begin{array}{rrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr$	pr	-0.026	-0.073	0.005
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	intercept	1.157	0.107	0.084
$\begin{array}{cccccccccccccccccccccccccccccccccccc$			$d_0 = 0, s = 1$	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	age	-0.077	0.001	0.006
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$age^{2}/100$	0.002	0.000	0.000
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$age^{3}/100$	0.000	0.000	0.000
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	d^*	-0.140	-0.006	0.006
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	pr	-0.018	-0.025	0.003
$\begin{array}{cccc} & \frac{d_0=1,s=0}{0.043} & 0.029\\ age^2/100 & 0.001 & 0.000 & 0.001\\ age^3/100 & 0.000 & 0.000 & 0.000\\ pr & -0.130 & -0.080 & 0.022\\ \text{intercept} & 1.205 & -0.261 & 0.338\\ \\ age & 0.043 & 0.034 & 0.031\\ age^2/100 & -0.001 & 0.000 & 0.001 \\ \end{array}$	intercept	1.723	0.750	0.074
$\begin{array}{cccccccccccccccccccccccccccccccccccc$				
age -0.043 0.042 0.029 $age^2/100$ 0.001 0.000 0.001 $age^3/100$ 0.000 0.000 0.000 pr -0.130 -0.080 0.022 intercept 1.205 -0.261 0.338 age 0.043 0.034 0.031 $age^2/100$ -0.001 0.000 0.001			$\underline{d_0 = 1, s = 0}$	
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	age	-0.043	0.042	0.029
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	$age^{2}/100$	0.001	0.000	0.001
$\begin{array}{ccccccc} pr & -0.130 & -0.080 & 0.022 \\ \text{intercept} & 1.205 & -0.261 & 0.338 \\ \\ age & 0.043 & \hline 0.034 & 0.031 \\ age^2/100 & -0.001 & 0.000 & 0.001 \\ \end{array}$	$age^{3}/100$	0.000	0.000	0.000
intercept 1.205 -0.261 0.338 $age 0.043 \frac{d_0 = 1, s = 1}{0.034} 0.031$ $age^2/100 -0.001 0.000 0.001$	pr	-0.130	-0.080	0.022
$\begin{array}{ccc} age & 0.043 & \underline{d_0=1,s=1}\\ age^2/100 & -0.001 & 0.000 & 0.001 \end{array}$	intercept	1.205	-0.261	0.338
$\begin{array}{cccccccccccccccccccccccccccccccccccc$			$d_0 = 1, s = 1$	
$age^2/100$ -0.001 0.000 0.001	aae	0.043	0.034	0.031
	$age^{2}/100$	-0.001	0.000	0.001
$age^3/100 0.000 0.000 0.000$	$age^{3}/100$	0.000	0.000	0.000
pr -0.109 -0.033 0.016	nr	-0.109	-0.033	0.016
intercept 0.338 -0.137 0.383	intercept	0.338	-0.137	0.383
		0.000		

Table 14: OLS Regression Coefficients: Employment

Moment	Data	Simulation	Standard Error
Mean of Ann	ual Earnings O	ver All Years	
$E(W d_0 = 0, s = 0)$	32300.000	36151.366	100.000
$E(W d_0 = 0, s = 1)$	50900.000	51881.700	100.000
$E(W d_0 = 1, s = 0)$	26000.000	29658.624	500.000
$E(W d_0 = 1, s = 1)$	40400.000	47926.220	600.000
	·	V CT 1	
Mean of Annual Earnings	in First Three	Years of Lab	our Market
$E(LnW s = 0, d_0 = 0, 1 \le t \le 3)$	9.390	9.406	0.018
$E(LnW s = 1, d_0 = 0, 4 \le t \le 6)$	9.700	9.712	0.012
$E(LnW s = 0, d_0 = 1, 1 \le t \le 3)$	9.200	9.167	0.052
$E(LnW s = 1, d_0 = 1, 4 \le t \le 6)$	9.530	9.606	0.047
Variano	e of Initial Ear	nings	
$Var(LnW _{\mathfrak{S}} = 0, d_{\mathfrak{S}} = 0, 1 = \frac{1}{\langle t \langle \mathfrak{S} \rangle}$) 0.486	0.032	0.015
$v u (D u v 3 = 0, u) = 0, 1 \leq t \leq 3$			
$Var(LnW s = 0, a_0 = 0, 1 \le t \le 0$ $Var(LnW s = 1, d_0 = 0, 4 \le t \le 0$	0.526	0.033	0.010
$Var(LnW s = 0, d_0 = 0, 1 \le t \le 0$ $Var(LnW s = 1, d_0 = 0, 4 \le t \le 6$ $Var(LnW s = 0, d_0 = 1, 1 \le t \le 3$	0.526 0.532	$0.033 \\ 0.316$	$\begin{array}{c} 0.010\\ 0.036\end{array}$

Table 15: Annual Earnings Distribution

Moment	Data	Simulation	Standard Error
		$d_0 = 0, s = 0$	
Q10	9.148	9.518	0.014
Q25	9.857	9.833	0.009
Q50	10.389	10.253	0.005
Q75	10.771	10.708	0.004
$\mathbf{Q90}$	11.060	11.126	0.004
		$d_0 = 0, s = 1$	
Q10	9.525	9.805	0.011
Q25	10.240	10.185	0.006
$\mathbf{Q50}$	10.751	10.652	0.003
Q75	11.126	11.115	0.003
$\mathbf{Q90}$	11.416	11.509	0.004
		$d_0 = 1, s = 0$	
Q10	8.556	9.193	0.072
Q25	9.278	9.638	0.044
Q50	9.971	10.035	0.030
Q75	10.454	10.488	0.022
Q90	10.919	10.950	0.032
		$d_0 = 1, s = 1$	
Q10	9.127	9.751	0.062
Q25	9.868	10.103	0.033
Q50	10.494	10.552	0.023
Q75	10.910	11.034	0.017
Q90	11.242	11.451	0.018

Table 16: Annual Earnings Quantiles

Moment	Data	Simulation	Standard Error
	2 0.00		
		$d_0 = 0, s = 0$	
age	0.123	0.155	0.003
$age^{2}/100$	-0.001	-0.002	0.000
d^*	-0.129	0.006	0.017
Intercept	7.625	7.051	0.051
		$\underline{d_0 = 0, s = 1}$	
age	0.172	0.214	0.002
$age^{2}/100$	-0.002	-0.002	0.000
d^*	-0.120	-0.001	0.011
Intercept	6.859	6.033	0.044
		$\underline{d_0 = 1, s = 0}$	
age	0.096	0.086	0.011
$age^{2}/100$	-0.001	-0.001	0.000
Intercept	7.713	7.858	0.197
		$\underline{d_0 = 1, s = 1}$	
age	0.190	0.197	0.012
$age^{2}/100$	-0.002	-0.002	0.000
Intercept	6.331	6.254	0.221

=

Table 17: OLS Regression Coefficients: Annual Earnings

Moment	Data	Simulation	Standard Error
	$d_0 = 0$	s = 0	
potential experience	$\frac{a_0}{0.110}$	0.112	0.006
potential experience ² /100	-0.204	-0.213	0.012
$E(v s=0, d_0=0)$	9.037	9.156	0.017
$Var(v s=0, d_0=0)$	0.345	0.192	0.014
$Var(\xi s = 0, d_0 = 0)$	0.209	0.099	0.004
$Cov(\epsilon_t, \epsilon_{t-1} s=0, d_0=0)$	0.087	0.085	0.002
$Cov(\epsilon_t, \epsilon_{t-2} s=0, d_0=0)$	0.053	0.075	0.002
	$d_0 = 0, s$	s = 1	
potential experience	0.145	0.126	0.004
potential experience ² /100	-0.285	-0.244	0.010
$E(v s=1, d_0=0)$	9.211	9.447	0.013
$Var(v s=1, d_0=0)$	0.364	0.181	0.014
$Var(\xi s = 1, d_0 = 0)$	0.230	0.090	0.003
$Cov(\epsilon_t, \epsilon_{t-1} s = 1, d_0 = 1)$	0.109	0.078	0.002
$Cov(\epsilon_t, \epsilon_{t-2} s=1, d_0=1)$	0.067	0.068	0.001
	$d_0 = 1, s$	s = 0	
potential experience	0.104	0.082	0.027
potential experience ² /100	-0.197	-0.142	0.064
$E(v s=0, d_0=1)$	8.761	8.830	0.061
$Var(v s=0, d_0=1)$	0.471	0.398	0.053
$Var(\xi s = 0, d_0 = 1)$	0.255	0.095	0.017
$Cov(\epsilon_t, \epsilon_{t-1} s=0, d_0=0)$	0.098	0.073	0.011
$Cov(\epsilon_t, \epsilon_{t-2} s=0, d_0=0)$	0.074	0.063	0.012
	$d_0 = 1, s$	s = 1	
potential experience	0.133	0.144	0.018
potential experience ^{2} /100	-0.323	-0.295	0.036
$E(v s=1, , d_0=1)$	9.245	9.158	0.053
$Var(v s=1, d_0=1)$	0.388	0.313	0.044
$Var(\xi s=1, d_0=1)$	0.263	0.061	0.015
$Cov(\epsilon_t, \epsilon_{t-1} s=1, d_0=1)$	0.127	0.050	0.009
	~ ~		

Table 18: First-Difference Regression on Annual Earnings

Moment	Data	Simulation	Standard Error
		$d_0 = 0, s = 0$	
Q10	8.182	8.609	0.025
Q25	8.655	8.834	0.016
Q50	9.094	9.116	0.014
Q75	9.457	9.446	0.012
$\mathbf{Q90}$	9.776	9.756	0.016
		$d_0 = 0, s = 1$	
Q10	8.390	8.892	0.019
Q25	8.816	9.145	0.012
Q50	9.236	9.443	0.010
Q75	9.579	9.763	0.010
$\mathbf{Q90}$	9.892	10.022	0.012
		$d_0 = 1, s = 0$	
Q10	7.953	8.033	0.084
Q25	8.396	8.423	0.049
$\mathbf{Q50}$	8.825	8.816	0.045
Q75	9.329	9.232	0.054
$\mathbf{Q90}$	9.716	9.648	0.049
		$d_0 = 1, s = 1$	
Q10	8.386	8.438	0.085
Q25	8.861	8.760	0.057
$\mathbf{Q50}$	9.292	9.149	0.059
Q75	9.700	9.565	0.043
Q90	10.002	9.894	0.080

Table 19: Fixed Effect Quantiles

References

- Baker, M., Stabile, M., and Deri, C. (2004). What do self-reported, objective, measures of health measure? Journal of human Resources, 39(4):1067–1093.
- Benítez-Silva, H., Buchinsky, M., Man Chan, H., Cheidvasser, S., and Rust, J. (2004). How large is the bias in self-reported disability? *Journal of Applied Econometrics*, 19(6):649–670.
- Black, N., Johnston, D. W., Shields, M. A., and Suziedelyte, A. (2017). Who provides inconsistent reports of their health status? the importance of age, cognitive ability and socioeconomic status. *Social Science* and *Medicine*, 191:9–18.
- Bound, J. (1989). The health and earnings of rejected disability insurance applicants. *The American Economic Review*, pages 482–503.
- Bound, J. and Burkhauser, R. V. (1999). Economic analysis of transfer programs targeted on people with disabilities. *Handbook of Labor Economics*, 3:3417–3528.
- Bruins, M., Duffy, J. A., Keane, M. P., and Smith Jr, A. A. (2018). Generalized indirect inference for discrete choice models. *Journal of Econometrics*, 205(1):177–203.
- Burkhauser, R. V., Daly, M. C., Houtenville, A. J., and Nargis, N. (2002). Self-reported work-limitation data: What they can and cannot tell us. *Demography*, 39(3):541–555.
- Dixon, K. A., Kruse, D., and Van Horn, C. E. (2003). Restricted access: A survey of employers about people with disabilities and lowering barriers to work. John J. Heldrich Center for Workforce Development, Rutgers.
- Grondin, C. (2016). A new survey measure of disability: The Disability Screening Questions (DSQ). Statistics Canada.
- Hartigan, J. A. and Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1):100–108.
- Keane, M. and Smith, A. A. (2003). Generalized indirect inference for discrete choice models. Yale University.
- Milligan, K. (2016). Canadian tax and credit simulator. Database, software and documentation, Version 2016-2.
- Office of the Auditor General of Canada (2015). Report 6-canada pension plan disability program. https://www.oag-bvg.gc.ca/internet/english/parl_oag_201602_06_e_41063.html.
- Shorrocks, A. F. et al. (2013). Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *Journal of Economic Inequality*, 11(1):99–126.
- Stern, S. (1989). Measuring the effect of disability on labor force participation. Journal of Human Resources, pages 361–395.