

Measuring the Value of Disability Insurance from Take-Up Decisions*

Andreas Haller, Norwegian School of Economics
Stefan Staubli, University of Calgary and NBER

July 2024

Abstract

This paper develops a novel sufficient statistics approach to identify the insurance value of disability insurance (DI) benefits. Our approach estimates the insurance value from the relative DI take-up responses to more generous DI benefits and the take-up responses to wage reductions. Empirically, we find that increasing Canadian DI benefits by \$1 creates a societal benefit of \$2.2 (insurance value). At the same time, we estimate that a \$1 increase in DI benefits costs society \$1.6. We thus find that the improved insurance outweighs the additional costs of higher benefits in the Canadian context.

Keywords: Disability insurance, work incentives, benefits, policy reform, sufficient statistics

JEL Codes: H53; H55; J14; J21; J65.

*Address: Andreas Haller, Norwegian School of Economics, 5045 Bergen, Norway; email andreas.haller@nhh.no. Haller is also affiliated with Cesifo and IZA. Stefan Staubli, University of Calgary, Calgary, AB T2N 1N4, Canada; email sstaubli@ucalgary.ca. Staubli is also affiliated with CEPR, IZA, and RSI Montréal. We thank Itzik Fadlon, Giulia Giupponi, Audrey Guo, Lucija Muehlenbachs, Adam Lavecchia, Attila Lindner, Todd Morris, Matthew Notowidigdo, Johannes Spinnewijn, Michael Veall, Josef Zweimüller and seminar participants at HEC Montréal, IFAU, McMaster University, University of Geneva, University of Munich, University of Queensland, the Bocconi/CEPR Social Security workshop, the National Tax Association conference, CESifo Area Conference on Public Economics, and the Nordic Public Policy Symposium. We thank Sencer Karademir and Kangyu Qiu for excellent research assistance. This research was supported by funding from the Social Sciences and Humanities Research Council in Canada. The analysis was conducted at the Prairie Regional RDC, as part of the Canadian Research Data Centre Network (CRDCN). This service is provided through the support of the University of Calgary, the province of Alberta, the Canada Foundation for Innovation, the Canadian Institutes of Health Research, the Social Sciences and Humanities Research Council, and Statistics Canada. All views expressed in this work are our own. All remaining errors are our own.

1 Introduction

Disability insurance (DI) programs provide valuable income replacement against losses in work potential due to a disability. At the same time, there are concerns that DI programs distort work incentives and reduce workers' labor supply. While a large and growing body of literature analyzes the impacts of DI on labor supply, we know little about the value of the program to DI beneficiaries. Yet, without knowledge of both the insurance value and incentive costs, assessing the welfare effects of a change in DI generosity remains elusive.

In this paper, we develop and implement a novel method to estimate the value of disability insurance. We show that the insurance value of disability benefits is identified by the ratio of the take-up responses to a change in disability benefit generosity and the take-up responses to a change in wages. The intuition behind this result is that the individual's DI take-up decision compares the utility when receiving DI benefits against the utility when working. Therefore, when a small change in DI benefits leads to a sizable increase in DI take-up, it suggests that the utility of receiving DI benefits substantially increased because of the higher DI benefits, i.e., the marginal utility of consumption when on DI is high.

Conversely, if a minor wage increase triggers a substantial reduction in DI take-up, it suggests that the attractiveness of remaining in the workforce has risen significantly, pointing to a high marginal utility of consumption when working. The relative DI take-up responses to higher benefits and lower wages therefore quantify the ratio of marginal utilities of consumption when on DI and when working, which captures the insurance value of DI benefits. Moreover, the DI take-up response to more generous benefits directly identifies the incentive costs of DI benefits. DI take-up responses are therefore sufficient statistics to evaluate the incentive-insurance trade-off in the DI program and can shed light on whether DI benefits are overly generous.

The traditional approach to estimating the value of social insurance programs focuses on consumption smoothing. The value of a change in social insurance generosity can be expressed by the change in consumption times workers' risk aversion (Gruber, 1997). Implementing this traditional approach is challenging because consumption

data are scarce, and the implied insurance value is sensitive to the level of risk aversion, which varies by context (Chetty and Szeidl, 2007). In contrast, our method does not require taking a stand on the right level of risk aversion and accommodates state-dependent utility functions. The main advantage of our approach is its broad applicability in settings where data on take-up decisions are available. The literature already estimates both the take-up effects of higher DI benefits and lower wages (e.g., Autor and Duggan, 2003; Milligan and Schirle, 2019). This literature has mostly studied how these push (weak labor markets) and pull (more generous DI) factors have contributed to the growth in the DI rolls. Generally, the literature has interpreted strong take-up responses to either increased benefits or negative economic shocks as a sign of moral hazard limiting the effectiveness of the DI program. We offer a new interpretation. A large take-up response to more generous benefits is not necessarily bad news. Instead, a strong response to more generous benefits, relative to the take-up response to reduced wages, indicates a high value of the DI program.

There are a couple of nuances to our approach. First, our approach identifies the insurance value of the marginal applicants. In the classic DI model of Diamond and Sheshinski (1995), the insurance value of the marginal applicant is representative for the program's overall insurance value. However, in models with richer heterogeneity and in models where the severity of disability does not enter the utility function additively the marginal type might not be representative. We argue that the marginal applicant's insurance value provides a lower bound for the program's insurance value. Moreover, the marginal applicant's insurance value is of interest on its own. Second, it is empirically challenging to isolate pure wage effects. The local economic shocks we exploit in our empirical implementation (Bartik instruments) arguably affect wages and employment. In a model extension we show that exploiting variation in earnings stemming from shocks, which affect both the wages and the employment margin, yields again a lower bound for the insurance value. The third nuance relates to the timing of effects. Empirically we exploit permanent changes in DI benefits and temporary economic shocks. We show that our results generalize to a dynamic model. In case of permanent benefit changes and temporary wage changes we can identify the insurance

value from take-up responses by rescaling the DI take-up response of the temporary earnings shock by its impact on the present value of lifetime earnings.

We implement our method in the context of Canada, allowing us to combine detailed administrative data from 20 percent of the Canadian population with exogenous variations in disability benefits and wages. We start our analysis by studying the DI take-up response to a 1987 reform, increasing disability benefits by 36 percent in the Canadian Pension Plan disability program (CPP-D) to align them with the level of disability benefits in the Quebec Pension Plan disability program (QPP-D). The QPP-D covers residents in the province of Quebec and the CPP-D covers residents in the rest of Canada. The differential variation in disability benefits in the CPP-D and QPP-D over time enables us to apply a differences-in-differences estimation approach. Our analysis shows that more generous disability benefits lead to more disability entry. The CPP-D disability take-up rate rises sharply after benefits become more generous. The increase is also persistent over time. Relating the take-up response to the change in benefits, we estimate a disability take-up elasticity of 0.58 (or a \$1000 increase in life-time DI benefits increases DI take-up by around 0.2 percentage points).

To estimate the impact of wages on disability take-up, we use a Bartik shift-share design that exploits variation in exposure to economic shocks driven by differential industry composition across Canadian census divisions. For robustness, we also exploit variation in labor market conditions across census divisions created by shocks to world oil and gas prices. Census divisions with a high employment share in the oil and gas sector are significantly more affected by oil and gas price shocks compared to census divisions with few workers in the oil and gas sector. Both approaches have been implemented in the literature to estimate DI take-up effects (Black et al., 2002; Autor and Duggan, 2003; Charles et al., 2018; Milligan and Schirle, 2019). We find that adverse economic shocks significantly increase DI take-up in line with the existing literature. In particular, we estimate that \$1000 reduction in life-time earnings increases DI take-up by 0.06 to 0.1 percentage points.

Combining the estimated take-up responses to wages and DI benefits, we find that the insurance value of a \$1 increase in DI benefits is \$2.2 in the Bartik shift-share design.

We estimate even higher insurance values of \$2.8 to \$3.4 for the oil price shocks. Our estimates thus suggest that the marginal utility gain from an additional dollar in the disabled state is more than twice as large as the utility gain from an extra dollar in the non-disabled state. At the same time, we find that incentive costs from providing more generous disability benefits are sizable but smaller. The fiscal cost of a \$1 increase in DI benefits is \$1.6, about three-quarters of the insurance value, suggesting that Canadian DI benefits are valuable and not too generous.

Applying our method to the U.S. estimates from Milligan and Schirle (2019) suggests that providing one additional dollar in DI benefits comes at a high cost of \$2.2. However, the implied insurance value is even higher with \$3.4.¹ Our result that the insurance value of DI benefits is high and exceeds the incentive costs is in line with recent papers that estimate the insurance value of the U.S. DI program. Deshpande and Lockwood (2022) estimate that the value of U.S. disability benefits exceeds a cost-equivalent tax cut by 64%. Cabral and Cullen (2019) infer from purchases of supplemental private insurance that the value of compulsory public DI is more than 2.5 times the cost of providing public DI. Using a structural model, Low and Pistaferri (2015) also find that welfare increases with increased DI generosity in the U.S. context, implying that the insurance value exceeds the incentive costs of DI benefits. Meyer and Mok (2019) estimate a consumption drop of 18% for U.S. DI recipients. The consumption drop multiplied by the coefficient of relative risk aversion plus one yields the insurance value (Gruber, 1997). Hence, the coefficient of relative risk aversion would need to be relatively high (at least 6) to match our Canadian estimate of the DI insurance value of 2.2, and very high (at least 13) to match the implied U.S. insurance value of 3.4.

Next to the above mentioned U.S. papers, Seibold et al. (2022) use the demand for private DI to infer the value of DI building on Einav et al. (2010) and Einav and Finkelstein (2011). They study a German reform that abolished a part of public DI and infer from limited take-up of private insurance that willingness-to-pay for DI in Germany is relatively low (the abolished DI coverage is valued at only 74% of the cost of providing the insurance). Haller et al. (2024) quantify the relative insurance

¹We explain in Section 6 how we use Milligan and Schirle (2019)'s estimates to construct the insurance value and incentive cost estimates.

losses from stricter DI eligibility criteria and lower DI benefits in Austria by exploiting spousal labor supply responses following the approach from Fadlon and Nielsen (2019). However, they do not quantify the absolute insurance value of DI benefits directly. While existing estimates of the insurance value of DI are based on spousal labor supply responses, private insurance purchases, consumption drops and structural models, the role of take-up decisions for the insurance value has not been recognized despite the large empirical literature on take-up.

The empirical literature on DI benefit generosity focuses primarily on how changes in benefit levels affect labor supply (for a review of the earlier literature, see Bound and Burkhauser, 1999). Most closely related to our study is a study by Gruber (2000) that exploits the same DI benefit reform in Canada. Gruber finds that a 36% increase in DI benefits induced an 11.5% increase in non-employment among 45-59-year-old men in the first two years after the reform. Instead, our paper estimates the impact of this reform on DI take-up, which is vital for inferring the insurance value of DI benefits. Moreover, we estimate impacts for the entire working-age population in the short and long-run, up to 13 years after the reform.

Our approach to estimating the impact of wage changes on DI take-up builds on several well-known US studies. Of particular relevance is the study by Autor and Duggan (2003), which uses a similar industry shift-share design to study the impact of labor market shocks across US states on DI take-up. Milligan and Schirle (2019) apply a similar shift-share design in both the Canadian and U.S. context. Because of data limitations, Milligan and Schirle (2019) use only three industries and five regions in their Canadian specification, potentially missing variation in local labor market conditions. We extend their analysis by exploiting variation across 260 census divisions and 102 industries. A critical contribution of our study is to clarify how temporary labor market shocks can be used to identify the DI take-up response to a permanent wage change. Another set of US studies estimates the impact of wage changes on DI take-up using booms and busts in coal mining (Black et al., 2002) and the oil and gas production (Charles et al., 2018). Since Canada is a large oil and gas producer, we use oil and gas price shocks as a robustness check and find quantitatively similar estimates as Charles

et al. (2018).

Methodologically, our paper contributes to the sufficient statistics literature that identifies the insurance value of social insurance benefits from observed behavior. Recent studies have developed methods to identify the insurance value of unemployment insurance (Landais and Spinnewijn, 2021; Hendren, 2017; Landais, 2015; Chetty, 2008, 2006; Gruber, 1997). Conceptually, our approach is related to the study by Chetty (2008), showing that the insurance value of UI benefits can be identified as the ratio of job search effort responses to changes in unconditional cash transfers and changes in UI benefits. Chetty (2008) emphasizes the exit margin (job search responses) to identify the value of UI benefits. We show that take-up responses reveal the insurance value of social insurance programs where the entry margin is important, such as in DI or old-age pensions. Our method can therefore be applied to other social insurance programs and only requires data on take up-decisions, benefits, and wages, which are more widely available in many countries.

The paper is organized as follows. The next section presents a model of disability insurance and derives formulas for optimal disability eligibility and benefits. Section 3 describes the data and the institutional background. Sections 4 and 5 present the empirical results on changes in DI benefit levels and changes in wages. Section 6 estimates the insurance value and the incentive costs of more generous DI benefits. Section 7 concludes.

2 Model: The Value of DI Benefits

To illustrate that take-up decisions identify the insurance value of DI benefits, we start with a simple version of the seminal DI model by Diamond and Sheshinski (1995). In the simple model, individuals work or receive DI benefits and only differ in their disability levels, which enters utility additively. In this model, the relative take-up responses of higher DI benefits vs. lower wages identify the insurance value of DI benefits.

We then extend the simple framework in three ways. We add a third labor market state (other benefits) and endogenous job search, we allow for non-additively separable

utility functions and richer heterogeneity beyond disability levels, and we provide a dynamic version of the model. With non-separability or heterogeneity beyond disability levels, the relative take-up decisions identify the insurance value of marginal applicants. The key question then is how representative the marginal applicant's insurance value is for the average DI recipient's insurance value. We argue that the marginal applicant's insurance value provides a lower bound on the average insurance value.

Setup. Consider a continuum of agents living for one period. Agents differ only in their level of disability θ , modeled as a random draw from a continuous distribution $F(\theta)$. An agent can decide to work or to apply for DI. An agent with disability level θ enjoys utility $u(w - \tau) - \theta$ if she works and utility $v(b)$ if she receives DI benefits. The outcome of a DI application is uncertain; an application is accepted with probability $p(\theta)$, where $p'(\theta) > 0$. Assuming no application costs, an agent applies if she prefers receiving DI benefits over working (i.e. if her disability level θ is high).² The “marginal applicant,” the agent who is indifferent between applying for DI benefits and remaining employed, has disability

$$\theta^A = u(w - \tau) - v(b). \quad (1)$$

Agents with disability $\theta \geq \theta^A$ apply for DI, while agents with disability $\theta < \theta^A$ remain employed.³

Welfare. The government maximizes a utilitarian social welfare function subject to a government budget constraint:

$$W(b) = V(b) + \lambda [G(b)], \quad (2)$$

²The assumption of no application cost is not critical for our result. With application costs ψ , the marginal applicant is determined by $p(\theta^A) (v(b) - [u(w - \tau) - \theta^A]) - \psi = 0$ and we can use the same approach to measure the insurance value of DI.

³The original framework of Diamond and Sheshinski (1995) includes a third labor market state “other welfare benefits”. To simplify notation, we do not consider this labor market state in the simple model without loss of generality. Our approach exploits the marginal applicant's responses and the marginal applicant is also determined by equation (1) in Diamond and Sheshinski (1995)'s framework.

where $V(b)$ denotes the aggregate indirect utility function and $G(b)$ denotes the total fiscal revenue. $V(b)$ is given by

$$V(b) = \int_0^{\theta^A} u(w - \tau) - \theta dF(\theta) + \int_{\theta^A}^{\infty} p(\theta)v(b) + (1 - p(\theta))(u(w - \tau) - \theta)dF(\theta) \quad (3)$$

and sums up the individual utilities. The first term measures the utility of individuals who do not apply to DI and work. The second term sums up the utilities of accepted DI applicants, $p(\theta)v(b)$, and of rejected DI applicants, $(1 - p(\theta))(u(w - \tau) - \theta)$. The total fiscal revenue $G(b)$ is the tax revenue from working individuals minus the DI program expenditures

$$G(b) = \tau[F(\theta^A) + \int_{\theta^A}^{\infty} (1 - p(\theta))dF(\theta)] - b \int_{\theta^A}^{\infty} p(\theta)dF(\theta), \quad (4)$$

which can be written as

$$G(b) = (1 - DI) \cdot \tau - DI \cdot b \quad (5)$$

where

$$DI \equiv \int_{\theta^A}^{\infty} p(\theta)dF(\theta). \quad (6)$$

Effect of Change in DI Benefits. An increase in the level of DI benefits has the following welfare effect

$$\frac{\partial W}{\partial b} = DI \cdot v'(b) - \lambda \left[DI + \frac{\partial DI}{\partial b} (b + \tau) \right]. \quad (7)$$

On the one hand, a one-dollar higher DI benefit level creates a direct welfare gain of size $DI \cdot v'(b)$ —all DI recipients benefit from the higher payments and value the extra money at their marginal utility of consumption. On the other hand, DI program costs mechanically increase by DI -dollars, and potential DI take-up responses create additional $\frac{\partial DI}{\partial b} (b + \tau)$ -dollars of expenditures. The optimal level of DI benefit trades off the more generous insurance against the higher program costs. The optimal DI benefit

therefore solves $\frac{\partial W}{\partial b} = 0$, which can be written as

$$\underbrace{\frac{v'(b)}{u'(w - \tau)}}_{\text{Insurance Value}} = 1 + \underbrace{\frac{\varepsilon_{DI,b}}{1 - DI}}_{\text{Multiplier}}, \quad (8)$$

where $\varepsilon_{DI,b} \equiv \frac{\partial DI}{\partial b} \frac{b}{DI}$ is the elasticity of DI take-up with respect to DI benefits.⁴ Formula (8) corresponds to the classic Baily-Chetty formula for the optimal level of social insurance benefits. The LHS of (8) measures the insurance value of DI benefits—the relative valuation of \$1 in the working state versus \$1 when on disability benefits. The RHS measures the total cost of a \$1 transfer between the employment and disability benefit state.

Inflow Responses Reveal Insurance Value of DI. We now explore how the insurance value can be identified from the relative DI take-up responses to higher DI benefits versus lower wages. The DI take-up effect of higher DI benefits is

$$\frac{\partial DI}{\partial b} = f(\theta^A) \cdot p(\theta^A) \cdot [v'(b)], \quad (9)$$

which follows from definition (6) and the application decision in (1). Similarly, the DI take-up effect of a reduction in wages is

$$-\frac{\partial DI}{\partial w} = f(\theta^A) \cdot p(\theta^A) \cdot [u'(w - \tau)]. \quad (10)$$

Together, the relative DI take-up response to benefits and wages (outside options) measures the insurance value:

$$-\frac{\partial DI}{\partial b} / \frac{\partial DI}{\partial w} = \frac{v'(b)}{u'(w - \tau)}. \quad (11)$$

Measuring the insurance value from take-up decisions is a new idea. The main advantage of this approach is that DI take-up effects can be estimated in many settings

⁴To rewrite (7) to (8) we use $\lambda = u'(w - \tau)$ and the balanced budget constraint (5) to substitute $\tau = \frac{DI}{1 - DI} \cdot b$.

with standard administrative data. Two influential strands of literature estimate the take-up effects of changes in DI benefit generosity and the DI take-up effects of labor market shocks (reduced wages). Our method combines these two strands of the empirical DI literature. Through the lens of our theory, we provide a novel interpretation of the relative take-up effects as a measure of the insurance value of DI benefits.

Intuition. Individuals apply to DI benefits when the utility they derive from being on DI surpasses the utility when working. When a slight adjustment in DI benefits leads to a sizable change in DI take-up, it indicates that a small benefit increase substantially boosts the utility derived from receiving DI benefits, i.e., the marginal utility of consumption when on DI is high. Conversely, if a minor wage increase significantly lowers DI take-up, it suggests that remaining in the workforce has become much more attractive, pointing to a high marginal utility of consumption when working. The relative DI take-up responses to higher benefits and lower wages therefore quantify the ratio of marginal utilities of consumption when on DI and when working (insurance value).

Another way to view this result is that DI benefits and wages are state-contingent transfers. One only receives DI benefits while on the DI program and wages only when working. If workers can perfectly self-insure against disability risk, they should be indifferent whether DI benefits or wages increase. The take-up responses to changes in benefits and wages are then the same because workers can transfer income across the states freely; the marginal utilities of consumption when on DI and when working are the same. In case of imperfect insurance the relative DI take-up responses identify the wedge in the ability to insure against disability risk.

How General Is This Result? A crucial question is whether our approach to identify the insurance value is limited to the specific model above or applies more generally. The model above is too simple in three main respects. First, it assumes that individuals only differ in their disability level and the disability level enters utility additively. Appendix A.1 extends the model to allow for non-additively separable utility functions and heterogeneity in wages, DI benefits, and unearned income. We show that with richer heterogeneity or non-separability our approach identifies the insurance value of

marginal applicants. We argue in Appendix A.1 that the insurance value of the marginal applicant provides a lower bound for the DI program’s average insurance value.

A second shortcoming of the above model is that it comprises only two labor market states (working or being on DI benefits). Having only two states is not a limitation from a theoretical perspective. As long as the marginal applicant returns to work if her DI application is rejected, (11) applies. Appendix A.2 provides a model with a third labor market state—other benefits—and endogenous job search to address an empirical challenge. We use local labor market shocks to identify pure wage effects, but these shocks could also affect separation and job finding rates. The extended model illustrates that if the shocks also affect the employment margin, the DI take-up response should be rescaled by the effect of the shocks on income because changes in income capture the joint impact on wages and employment. Moreover, we show that with employment responses the relative take-up responses provide a lower bound on the insurance value.

The third shortcoming of the simple model is that it is static. Our approach to estimating the insurance value generalizes to a dynamic setting, as Appendix A.3 demonstrates. The challenge is to connect the model to the empirical estimates. While the variation in DI benefits is permanent, the variation in wages is temporary. We show that we can identify the insurance value from permanent benefit changes and temporary wage changes if we rescale the DI take-up response of the temporary shock by its impact on the present value of lifetime earnings.

3 Institutional Background and Data

Our setting to estimate the value of disability insurance benefits is Canada. This section describes the Canadian public long-term DI program, the variation in DI benefits and earnings we exploit in the empirical analysis, and the data.⁵

⁵For more details on the Canadian DI program, see Torjman (2002), Baker and Milligan (2012), and Campolieti and Riddell (2012).

3.1 Institutional Background and Policy Variation

A peculiarity of Canada is the existence of two separate DI programs, both established in 1966 and financed by a payroll tax. Quebec, the second largest province with about a fourth of the Canadian population, desired to retain control over the pension plan's design and created its own program, the Quebec Pension Plan Disability (QPP-D). In contrast, the other provinces are all covered by the Canadian Pension Plan Disability (CPP-D) program. Today, the programs are nearly identical concerning eligibility criteria and level of benefits, but historically QPP-D benefits were higher until a 1987 reform increased the CPP-D benefits. We exploit this reform to estimate the impact of benefit generosity on DI take-up.

To qualify for DI benefits, an individual needs to suffer from a prolonged—lasting at least one year—mental or physical disability that prevents pursuing any substantially gainful employment. Eligibility also depends on the contribution time since age 18. Before 1987, individuals in both programs had to contribute for at least 5 out of the last 10 years or one-third of the contribution period. The CPP-D relaxed contribution requirements in 1987, allowing workers who contributed 2 of the past 3 years to also qualify. It tightened the requirements in 1997 (4 contribution years in the last 6 years) and relaxed them again in 2005 (3 contribution years in the last 6 years or 25 years in total). The QPP-D changed the contribution period only once. Since 1993 individuals are eligible if they contributed for at least half of the contribution period, 5 of the past 10 years, or 2 of the last 3 years. Once benefits are awarded, DI recipients receive monthly payments until they die, return to work, or reach the retirement age.

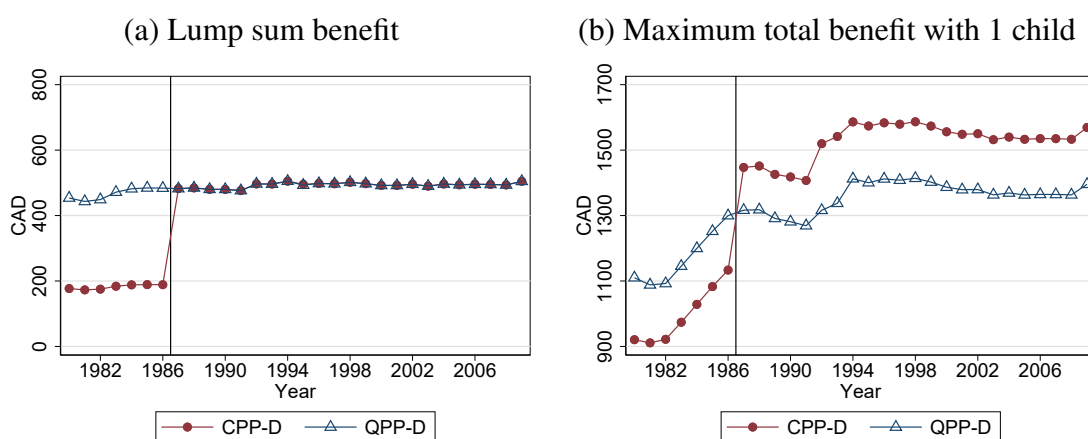
Variation in Potential DI Benefits. DI benefits in the CPP-D and the QPP-D consist of three parts: a lump-sum benefit identical for all eligible recipients, an earnings-related benefit, and a child allowance, which is a fixed amount per child under age 18. The earnings-related benefit is calculated in the same way in both programs, but after 1972 the lump-sum component grew much faster in the QPP-D compared to the CPP-D.⁶ By 1986, the lump-sum transfer in the QPP-D was almost three times as large as in

⁶For the earnings-related benefit the DI recipients earnings history is inflated by a wage index and the lowest 15 percent of monthly real earnings are dropped. The

the CPP-D, as illustrated in Panel (a) of Figure 1.

In an effort to align the two programs, in January 1987 the government raised the CPP-D lump-sum benefit to the same level as in the QPP-D. This change increased the annual CPP-D benefits by about CAD 3,500 (in 2019 dollars) or about 36 percent. Panel (b) of Figure 1 shows that before the reform the maximum CPP-D benefits were lower than in the QPP-D, but the 1987 reform raised CPP-D benefits above the level in the QPP-D. After the reform, the average monthly DI benefit payments moved in parallel, except for minor differences between 1992 and 1994. These differences arose because the CPP-D increased the child component in 1992, while the QPP-D followed only in 1994.

Figure 1: Maximum monthly DI benefit payments in CPPD and QPPD 1980-1992



Notes: The figure shows the maximum lump sum DI benefit (panel a) and maximum total DI benefit with one child (panel b) in the CPP-D and QPP-D (in 2019 Canadian Dollars). Before the reform in 1987 DI benefits in CPP-D were less generous than in QPP-D. The adjustment in the lump-sum component of the CPP-D benefits in the 1987 reform raised CPP-D benefits above the QPP-D level. Pre- and post-reform the CPP-D and QPP-D benefits evolve in parallel. Numbers are based on the CPP STATS BOOK 2019 and “Évolution de la clientèle de la rente d’invalidité de 1970 à 2010” (Diarra et al., 2015).

The 1987 reform implemented two additional changes to the CPP that could matter for our analysis. As discussed above, the first change was a relaxation in the contribution requirement, permitting workers with 2 contribution years in the last 3 years to qualify for benefits. To isolate the benefit generosity effect, we would ideally restrict the sample to individuals who fulfill pre- and post-reform eligibility criteria. But since our data start only in 1982, we cannot verify the pre-reform contribution requirement earnings-related benefit is then calculated as 18.75 percent of the average monthly earnings of the remaining earnings history.

over the last 10 years. Instead, we look at individuals in 1992 for whom we can observe the contribution history over the last 10 years. We find that only 2.3% would qualify with 2 out of 3 contribution years alone. Almost 87% would already qualify under the pre-reform contribution requirement, while 11% never qualify. Hence, the potential impact of the relaxation in contribution requirements is small. The 1987 reform also lowered the early retirement age in the CPP from age 65 to 60. We therefore restrict our sample to ages 15 to 59. Still, forward-looking individuals might adjust labor supply before age 60, but such anticipatory effects are likely small, as previous literature finds that the change had little impact on the labor supply of 60-64-year-olds (Baker and Benjamin, 1999; Staubli and Zhao, 2023). Moreover, the QPP lowered its early retirement age already in 1984, and we find that this change had no impact on labor supply or DI take-up before age 60.

Variation in Potential Earnings. To estimate the impact of potential earnings on DI take-up, we combine granular variation in industrial composition across Census Divisions (CDs) with national-level changes in employment to predict local shocks in potential earnings. Canada has about 260 CDs, which differ significantly in their industry composition. For example, manufacturing industries such as auto-making, food and beverage, and fabricated metals are concentrated in CDs in Ontario, while the technology sector is strong in Quebec CDs. The variation implies that industry-specific shocks will impact some CDs much more than others.

We leverage the cross-CD variation in industrial composition and national-level changes in employment using a Bartik shift and share instrumental variable approach. Our main specification exploits variation in industry composition across three-digit industries (102 industries) similar to Autor and Duggan (2003). We probe the robustness of our results using variation in oil and gas employment across CDs as in Charles et al. (2018), paired with national changes in oil and gas employment or prices. We use both national employment and prices changes, since Canadian oil prices were under price controls from 1974 to 1985.

Oil and gas employment is concentrated in CDs in the provinces of Alberta, Saskatchewan, British Columbia, and Newfoundland and Labrador. Appendix Figure

C.3 displays the real oil price trends between 1982 and 2016.

3.2 Data

For our analysis, we use individual tax return data from the Longitudinal Administrative Databank (LAD), a representative panel of 20 percent of Canadian tax filers between 1982 and 2019. Individuals who are selected into the LAD are followed each year they file a return and they can be linked across years using an anonymous identifier. The LAD contains information on earnings, government transfers—including Canada and Quebec pension plan benefits—taxes, and demographics. The data also contain detailed geographic information including the census division and province of residence, which we use to infer whether an individual is covered by the CPP-D or QPP-D.

We measure DI benefit receipt by whether an individual receives a public pension. The data does not distinguish between the type of public pension benefits received, but the only available benefit before age 60 other than DI are survivor benefits. Few people claim survivor benefits before age 60 because widowhood before age 60 is rare. The 1987 reform also did not change the generosity of survivor benefits. We would thus not expect any change in the take-up of survivor benefits around the reform. Since the data record the spouse's year of death, we can approximate the receipt of survivor benefits using spousal deaths. In a robustness check, we set DI benefit receipt to zero for individuals who start receiving benefits within a year their spouse died. The results are nearly identical.

One drawback of the LAD is that it does not record an individual's occupation or industry, which we need to calculate the CD industry shares. To address this issue, we pool data from the long-form Canadian census for the years 1981, 1986, 1991, 1996, 2001, 2006, 2011, and 2016. The long-form census contains the 3-digit industry of employment for 20 percent of the population.

Analysis Samples. To study the effect of more generous DI benefits, we focus on 15-59-year-old individuals who continuously file taxes from 1982 until 1992 or until they

die, whichever is first.⁷ We impose this sample restriction to have a balanced sample for the main period of analysis.

To study the effect of changes in potential earnings, we focus on 15-59-year-old individuals in the LAD from 1982 to 2016. We observe individuals in five-year intervals because the Canadian Census is only conducted every 5 years and we need the Census to calculate industry shares.⁸

4 Impact of Benefit Generosity

4.1 Estimation Strategy

The 1987 reform increased CPP-D benefits while leaving QPP-D benefits unchanged. We exploit this policy-induced variation in benefits in a difference-in-differences (DiD) design, comparing the change in an outcome variable in the Rest of Canada (RoC) with the change in the same outcome variable in Quebec over time. This comparison can be implemented with the following regression:

$$Y_{ipt} = \alpha + \sum_{s=1982, s \neq 1986}^{2000} \beta_s (I[p = RoC] \cdot I[s = t]) + \theta_p + \pi_t + X'_{ipt} \delta + \epsilon_{ipt}, \quad (12)$$

where Y_{ipt} is an outcome variable such as an indicator for DI take-up of individual i living in province p in year t , $I[p = RoC]$ is a dummy equal to 1 if an individual lives in the RoC, $I[t = s]$ is an indicator for the observation being in year s , θ_p are province fixed effects, π_t are year fixed effects, and X_{ipt} is a vector of demographic and labor market characteristics (e.g., age, gender, and the provincial unemployment rate). Gruber (2000) studies the effect of the 1987 reform on labor force non-participation for 45-59 year old men using the Canadian Survey of Consumer Finances from 1985-1989 and a “2x2” difference-in-difference design. We build on his work by considering the impact on DI take-up—a central sufficient statistic for welfare analysis—for the full

⁷Appendix Table B.1 shows summary statistics for this sample, separately for Quebec and the Rest of Canada before and after 1987.

⁸Since the LAD only starts in 1982, we combine Census industry shares in 1981 with LAD variables in 1982.

labor force population of 15-59 year old men and women and we estimate the dynamic effects in each year between 1982 and 2000.

The coefficients of interest are the β_s , each β_s -coefficient measures the average causal effect of the reform-induced benefit increase in year s relative to the base year, 1986. The pre-reform β_s -coefficients ($s < 1987$) provide pre-tests for spurious trends. They should not be statistically significant if the identification assumption of parallel trends holds, although they could pick up anticipation effects. Such effects are unlikely, as the reform was only enacted six months before it became effective. We cluster the standard errors at the census division level (roughly 260 clusters).

For our main effects of the reform, we use the 1991-coefficient (β_{1991}), which measures the causal impact of higher DI benefits on individuals in the RoC relative to those in Quebec in 1991. We stop in 1991 for two reasons. First, the 1991-coefficient measures the impact five years after the reform, matching the time window of our wage effects analysis, where we also look at five-year changes due to the census's five-year intervals. Second, in 1992, the CPP-D lifted the time limit on late applications. Initially, individuals who applied for CPP-D benefits more than 15 months after the onset of a disability were automatically denied. The 1992 change abolished the 15-month time limit.⁹ Moreover, the CPP-D increased the child component of DI benefits in 1992, and the QPP-D followed in 1994, leading to small differences in benefit generosity trends over these two years, as discussed in Section 3. Still, to assess the long-run impacts of more generous benefits, we also document the effects until the year 2000 but are more cautious in interpreting the estimates after 1992 as causal.

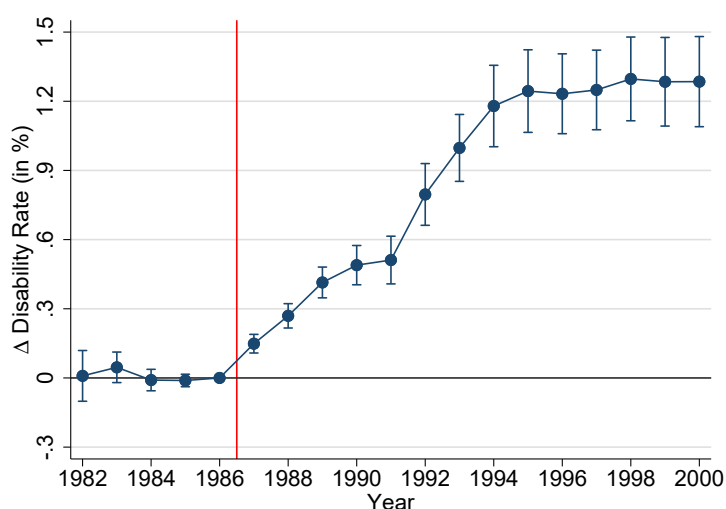
4.2 Empirical Results

We start our analysis by plotting the β_s -coefficients from regression (12). Figure 2 shows that higher DI benefits induce more entry into DI (Appendix Figure B.1 shows the raw trends in RoC and Quebec). Before the reform DI take-up does not differ between RoC and Quebec, providing evidence in support of parallel trends in DI take-

⁹See table 1 in Campolieti and Riddell (2012) for a nice overview of the policy variation after 1992.

up. When the reform becomes effective in 1987, DI take-up starts to increase in RoC relative to Quebec. Because of the reform DI take-up in RoC relative to Quebec is about 0.5 percentage points higher in 1991. By 1994, the disability rate is around 1.2 percentage points higher when the effect starts to level off. Appendix Figure B.2 shows the estimates of equation (12) when we set DI benefit receipt to zero if the spouse died within one year of benefit take-up to account for take-up of survivor benefits. The patterns and magnitudes of the estimates are almost identical.

Figure 2: Difference in DI Take-Up Between RoC and Quebec by Year



Notes: The figure shows the estimated β_s -coefficients from the econometric specification in (12). The capped spikes denote the upper and lower end of the 95-percent confidence interval.

Figure 3 plots the β_s -coefficients for the log of disability benefits in Panel (a), non-employment (an indicator for zero earnings on the tax return) in Panel (b), and the log of earnings in Panel (c) (Appendix Figure B.1 plots raw trends in RoC and Quebec). In line with the significant increase in DI take-up, log DI benefits continuously increase in RoC relative to Quebec after the reform and level off after 1994. Not being able to pursue substantial gainful employment is an essential eligibility requirement for DI benefits. Panel (b) shows that the reform reduced employment: the rate of non-employment rises more sharply in RoC compared to Quebec after the reform. The non-employment effect is almost twice as large as the DI take-up effect (but also less precisely estimated).¹⁰

¹⁰Absent individual data on applications, we tried to obtain aggregate application

Finally, Panel (c) shows that the reform lowered earnings, consistent with the rise in non-employment.

Figure 3: Difference in Other Outcomes Between RoC and Quebec by Year



Notes: The figure shows the estimated β_s -coefficients from the econometric specification in (12). The capped spikes denote the upper and lower end of the 95-percent confidence interval. Log disability benefits and log earnings are constructed as $\log(1 + x)$ where x are annual disability benefits or earnings.

Table 1 presents the reform effect five years after the reform (β_{1991} from regression (12)). More generous DI benefits raise DI benefit receipt by 0.527 percentage points and DI benefits by 0.047 percent. The estimate for DI benefit receipt implies an elasticity of DI benefit receipt with respect to DI benefits of 0.58. This elasticity identifies the fiscal externality in our model for optimal DI benefits, similar to the standard Bailey-Chetty model for unemployment insurance. It implies that a one-dollar increase in DI benefits data through a freedom of information request to estimate the impact of the reform on application behavior, but were informed that such data do not exist for the CPP-D.

Table 1: Impacts of 1987 Reform

| | DI benefit receipt | Log DI benefits | Non-employment | Log earnings |
|------------------------------------|---------------------|---------------------|--------------------|---------------------|
| Coeff. estimate (β_{1991}) | 0.527*** (0.061) | 0.047*** (0.006) | 1.689** (0.677) | 0.043*** (0.007) |
| Mean | 1.84 | 0.163 | 14.21 | 0.186 |
| Elasticity | 0.580*** (0.067) | 0.096*** (0.011) | 0.241** (0.096) | 0.469*** (0.080) |

Notes: This table reports the reform effect five years after the reform for different outcomes (the β_{1991} -coefficient from equation (12)). Standard errors are reported in parentheses and are clustered at the Census Division (CD). The sample include 15-59 year-old individuals. All specifications include year fixed effects, province fixed effects, age and its square, and the provincial unemployment rate. Levels of significance: *10%, **5%, and ***1%.

costs tax payers a total of 1.6 dollars because of induced entry into DI.

Non-employment increases by 1.68 percentage points and earnings fall by 0.043 percent. The implied elasticity of non-employment with respect to disability benefits is 0.24, which is consistent with Gruber (2000)'s findings, who reports non-participation elasticities of 0.28–0.36 for older male Canadian workers.

5 Impact of Wage Shocks

5.1 Estimation Strategy

To assess the effects of a change in earnings on disability take-up, we follow previous literature (Black et al. (2002); Autor and Duggan (2003); Charles et al. (2018)) and estimate first-difference regressions of the form:

$$\Delta y_{ict} = \alpha + \beta(\Delta E_{ict}) + \lambda_t + \delta \Delta X_{ict} + \varepsilon_{ict}, \quad (13)$$

where Δ denotes the first difference operator over five-year intervals. The primary outcome variable Δy_{ict} is the change in DI claiming for individual i living in census division c in year t , ΔE_{ict} is the change in annual earnings, λ_t is a vector of year dummy variables, ΔX_{ict} is a vector of changes in control variables (age and age-squared), and ε_{ict} are any unobserved factors affecting DI claiming such as tastes for work. The main parameter of interest is β , the effect of a change in potential earnings on DI claiming.

We expect that $\beta < 0$ as higher potential earnings should reduce the likelihood to claim DI benefits. A concern in estimating β using equation (13) is that ΔE_{ict} is endogenous as changes in earnings are likely correlated with changes in unobserved factors affecting DI claiming. For example, the OLS estimate of β will be biased downward if unobserved tastes for work, ε_{ict} , are positively correlated with earnings and negatively correlated with disability claiming. To surmount this endogeneity concern, we require plausibly exogenous changes to potential earnings. We follow Autor and Duggan (2003) and instrument for ΔE_{ict} using a Bartik shift-share design that exploits variation in industry composition across census division (CD) and national-level changes in employment to predict CD-specific employment growth:

$$\Delta \hat{\gamma}_{ct} = \sum \omega_{kct-5} \cdot \Delta \gamma_{\mathcal{K}t}, \quad (14)$$

where $\Delta \hat{\gamma}_{ct}$ is the predicted log employment change for each CD c between $t - 5$ and t , ω_{kct-5} is the share of employment in industry k in CD c in year $t - 5$, and $\Delta \gamma_{\mathcal{K}t}$ is the log change in the national employment share of the three-digit industry k . As in Autor and Duggan (2003), we exclude each CD's industry k employment when calculating $\Delta \gamma_{\mathcal{K}t}$, hence the subscript \mathcal{K} .

To probe the robustness of our results, we also exploit variation in oil and gas employment across CDs, paired with national-level changes in oil and gas employment and prices, similar to Charles et al. (2018). In this case, predicted employment growth is $\Delta \hat{\gamma}_{ct} = Oil_{ct-5} \cdot \Delta P_t$ where Oil_{ct-5} is the oil and gas employment share in CD c in year $t - 5$ and ΔP_t is the national change in either oil and gas employment or the national change in the price of oil between $t - 5$ and t . As in Black et al. (2002) and Charles et al. (2018), to account for sluggish local business adjustments following international price shocks, we include two lags of the instrument in ΔP when looking at oil prices.

Using local labor market shocks from variation in local industry composition creates three challenges for estimating the insurance value in equation (11). First, we would like to identify the pure wage effects on disability claiming, but local labor market shocks may also impact job separations and job findings. Appendix A.2 shows that if labor market shocks impact job separations and job findings, we get a lower bound of

the insurance value if we scale the DI inflow effect by disposable income—the sum of earnings, DI benefits, and unemployment insurance (UI) benefits.

Second, local labor market shocks create *temporary* variation in wages, while the 1987 reform induces a *permanent* increase in DI benefits. To compare the effects of a temporary versus a permanent change, Appendix A.3 extends the static model to a dynamic model. The model illustrates that lifetime income (instead of annual income) captures the permanent impact of a temporary wage shock. We measure average lifetime income in year t by summing up disposable income from year t until age 60 using a discount rate of 3 percent and dividing by the number of years between year t and age 60.

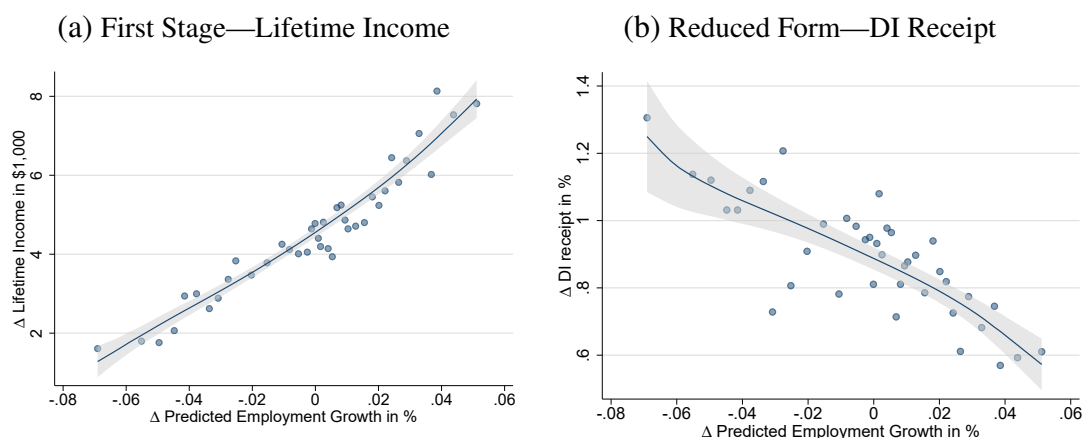
Third, our model assumes that marginal applicants for changes in benefits and changes in wages are comparable. Otherwise, differential take-up responses to benefits and wages could reflect a compositional effect and might not accurately reflect differences in marginal utility between states. The empirical concern is that local economic shocks may affect a different segment of the population compared to the segment affected by changes in DI benefits. To address this concern, we perform a complier analysis and compare the characteristics of the marginal types responding to local economic shocks and changes in DI benefits.

5.2 Empirical Results

Figure 4 provides a graphical representation of our estimation strategy. The left panel presents the visual first stage, obtained from a local linear regression of the change in lifetime income in \$1,000 against the predicted employment growth from the industry shift-share instrument (the shaded area indicates a 95 percent confidence interval). The change in lifetime income is monotonically increasing in the predicted employment growth and close to linear. The right panel plots the reduced-form effect of predicted employment growth on DI receipt using a local linear regression. DI receipt is monotonically decreasing in predicted employment growth.

Table 2 presents the OLS and 2SLS estimates of equation (13) for different earnings and income measures. Our OLS estimates show a negative impact of earnings and

Figure 4: Lifetime Earnings and DI Receipt by Predicted Employment Growth



Notes: Effects of Predicted Employment Growth Rates at the Census Division on Lifetime Income (Panel a) and DI Receipt (Panel b).

(lifetime) income on disability claiming, but the effects are quantitatively small consistent with findings by Autor and Duggan (2003) and Charles et al. (2018) for the US. The 2SLS estimates are significantly larger in magnitude. The second column shows that a \$1,000 increase in earnings reduces DI enrollment by 0.044 percentage points. We find a similar effect for current income (column 4). In contrast, the impact of a \$1,000 increase in lifetime income on DI claiming is about twice as large. This pattern is consistent with local labor market shocks having a large impact on current earnings and income, but a smaller impact on lifetime income because the shocks are temporary. Indeed, the first-stage coefficients is about half as large for lifetime income compared to current earnings and income. The heteroscedasticity-robust F-statistic (Montiel Olea and Pfluger, 2013) testing the significance of the excluded instruments in the first-stage equation is 50 or larger, well above conventional thresholds.¹¹

Appendix Table C.3 reports the 2SLS estimates and first-stage coefficients for the oil instruments. The instruments are strong and consistent with the shift-share estimates in Table 2 we find that higher earnings and income reduce DI take-up significantly. The magnitude of the effect is also consistent but the point estimates are slightly smaller: a \$1000 increase in lifetime income reduces DI take-up between 0.061 and 0.072 per-

¹¹Montiel Olea and Pfluger (2013) find that the critical values of the F-statistic to a test of IV relative bias not exceeding 10 percent with a significance level of 5 is between 11 and 23.1, analogous to the Stock and Yogo (2005) rule-of-thumb cutoff of 10.

Table 2: The Impact of Earnings/Income on Disability Insurance Enrollment

| | Earnings (in \$1,000) | | Current income (\$1,000) (in \$1,000) | | Lifetime income (in \$1,000) | |
|------------------------|--------------------------|----------------------|--|----------------------|---------------------------------|----------------------|
| | OLS | 2SLS | OLS | 2SLS | OLS | 2SLS |
| Δ DI enrollment | -0.004*** (0.001) | -0.044*** (0.008) | -0.007*** (0.002) | -0.044*** (0.008) | -0.015*** (0.003) | -0.096*** (0.018) |
| 1st-stage coefficient | | 117.0*** (15.6) | | 116.0*** (15.6) | | 53.0*** (7.7) |
| Effective F-statistic | | 55.9 | | 55.2 | | 47.3 |
| Obs. | 18,829,205 | | 18,829,205 | | 18,829,205 | |

Notes: This table reports the OLS and 2SLS estimates of equation (13) for different earnings and income measures. Standard errors are reported in parentheses and are clustered at the Census Division (CD). The sample include 15-59 year-old individuals between 1982 and 2016. Levels of significance: *10%, **5%, and ***1%.

centage points.

Complier Analysis. We perform a complier analysis to compare characteristics of marginal applicants for changes in benefits versus changes in wages. We cannot identify individual marginal applicants, but it is possible to describe their observable characteristics. Following Frandsen et al. (2023), who extend Abadie’s (2003) method for estimating complier characteristics to non-binary instruments, we run the following 2SLS specification in first differences:

$$(X \cdot \Delta DI)_{it} = \alpha + \beta \Delta DI_{it} + \pi_t + \varepsilon_{it}, \quad (15)$$

where X is an exogenous characteristic, ΔDI is a dummy for DI entry in year t and π_t are year dummies. For the benefit change, we instrument ΔDI with $(RoC_p \cdot A87_t)$, the interaction of a dummy for living in the Rest of Canada and a dummy for the year 1987 or greater. This interaction captures the quasi-experimental variation in DI benefits from the 1987 reform. For the wage change, we instrument ΔDI with the shift-share instrument $\Delta \hat{\gamma}_{ct}$. Frandsen et al. (2023) show that β yields a weighted average of characteristic X among marginal applicants who take up DI because of a benefit or wage change.

Table 3 reports estimates of equation (15) for different characteristics. The p-value in the last column tests the null hypothesis that marginal applicant characteristics are identical for a benefit and a wage change. The first two rows display the average age and the share of females among marginal applicants. The marginal applicants of benefit and wage changes are very similar in these two characteristics and the small differences are not statistically significant (p-values of 0.269 and 0.807). Marginal applicants are older with an average age of around 50 and more likely to be male (the female share is around 40%). The third row looks at the probability of having died by age 60, a proxy for disability severity θ . We find that marginal applicants' mortality rates by age 60 are comparable (13% for a benefit change and 12% for a wage change), and we cannot reject the null hypothesis that the coefficients are identical ($p=0.612$). Moreover, we find that the share of marginal applicants across different quartiles of the earnings and taxable income distribution is strikingly similar. Marginal applicants are over-represented in the lower earnings quartiles. Approximately 40% of marginal applicants come from the lowest quartile, 20-25% from the middle quartiles, and around 10% from the highest earnings quartile.

Taken together, these results indicate that our empirical variation in DI benefits and wages affect similar types and we thus identify meaningful differences in marginal utilities, and not composition effects.¹²

6 Welfare Effects

This section estimates the insurance value and welfare impacts of changes in DI benefits. For our Canadian implementation, we combine the estimates from Sections 4 and 5. Additionally, we provide a back-of-the-envelope implementation of our approach for the U.S. DI system using estimates from the literature.

¹²Appendix A.1 discusses this point more formally. Equation (A.3) illustrates that we want the same (w, A, b, θ) -marginal types for both benefits and wage changes. Table 3 indicates that the empirical variation in DI benefits and wages that we exploit affect similar (w, A, θ) -types. Since DI benefits are a deterministic function of past earnings, we have comparable (w, A, b, θ) -types for benefit and wages changes.

Table 3: Comparison of Complier Characteristics

| | Benefit change | | Wage change | | Equality test |
|----------------------------|----------------|-----------|-------------|-----------|---------------|
| | Coeff. | Std. err. | Coeff. | Std. err. | (p-value) |
| Age | 50.57*** | 0.80 | 53.12*** | 2.20 | 0.269 |
| Share female | 0.426*** | 0.038 | 0.409*** | 0.054 | 0.807 |
| Pr(died by age 60) | 0.132*** | 0.017 | 0.119*** | 0.019 | 0.612 |
| Share earnings 1. quartile | 0.377*** | 0.047 | 0.441*** | 0.058 | 0.392 |
| Share earnings 2. quartile | 0.274*** | 0.026 | 0.256*** | 0.050 | 0.754 |
| Share earnings 3. quartile | 0.213*** | 0.032 | 0.238*** | 0.034 | 0.598 |
| Share earnings 4. quartile | 0.136*** | 0.034 | 0.065 | 0.063 | 0.320 |
| Share income 1. quartile | 0.357*** | 0.047 | 0.345*** | 0.040 | 0.849 |
| Share income 2. quartile | 0.268*** | 0.030 | 0.292*** | 0.053 | 0.703 |
| Share income 3. quartile | 0.223*** | 0.025 | 0.264*** | 0.030 | 0.285 |
| Share income 4. quartile | 0.152*** | 0.033 | 0.099 | 0.062 | 0.450 |

Notes: This table reports the complier characteristics (β -estimates of equation (15)) for benefit and wage changes. The last column reports the p-value for the null hypothesis that marginal applicant characteristics are identical for benefit and wage changes. Levels of significance: *10%, **5%, and ***1%.

Canadian Implementation. Panel A of Table 4 reprints the DI take-up effects of a \$1,000 increase in DI benefits and a \$1,000 increase in lifetime income for the different instruments. The first column in Panel B measures the fiscal multiplier of increasing DI benefits by one dollar. The fiscal multiplier is the sum of the direct costs—an additional dollar to current beneficiaries— plus indirect costs from more people entering the DI program. The indirect cost is captured by the elasticity of DI take-up with respect to benefits, normalized by the share of people not on DI. We estimate that the fiscal multiplier of an additional dollar in DI benefits is 1.6 dollars.

The last three columns of Panel B report the insurance value of increasing DI benefits by one dollar, which we obtain by dividing the DI take-up response to a change in DI benefits by the take-up response to a change in lifetime income. We find that the insurance value of an additional dollar in benefits is 2.2 dollars when using the industry shift-share design. In other words, the utility gain from an extra dollar in the disabled state is 2.2 times larger than the marginal utility gain from an extra dollar in the non-disabled state. We find even higher insurance values of 2.8-3.4 for the oil price shocks.

Our point estimates imply that an increase in Canadian DI benefits would improve welfare. However, the insurance value is estimated with noise. We can reject perfect in-

Table 4: Welfare Analysis

| <i>A. Δ DI enrollment per \$1,000</i> | | | | |
|---|---|---|----------------------|----------------------|
| | DI benefits ($\partial DI/\partial b$) | Lifetime income ($\partial DI/\partial w$) | | |
| | | Industry share | Oil employment | Oil price |
| Coeff. estimate | 0.208*** (0.024) | -0.096*** (0.018) | -0.061*** (0.015) | -0.072*** (0.007) |
| <i>B. Welfare impacts</i> | | | | |
| | Multiplier ($1 + \frac{\varepsilon_{DI,b}}{1-DI}$) | Insurance value ($-\frac{\partial DI/\partial b}{\partial DI/\partial w}$) | | |
| | | Industry share | Oil employment | Oil price |
| Estimate | 1.591*** (0.069) | 2.166*** (0.476) | 3.421*** (0.923) | 2.885*** (0.438) |
| P-value: estimate = 1 | 0.000 | 0.014 | 0.009 | 0.000 |
| P-value: multiplier = ins. value | | 0.195 | 0.041 | 0.001 |

Notes: This table combines the estimates from Sections 4 and 5 to construct the multiplier and insurance value of higher DI benefits. Levels of significance: *10%, **5%, and ***1%.

surance (i.e., an insurance value of one) with high confidence, as indicated by p-values around and below 1% in Panel B of Table 4. When we test whether the insurance value equals the multiplier, we find p-values of 0.195 for the industry shift-share design, and 0.041 and 0.001 for the oil specifications. Hence, for the shift-share design, we cannot reject that Canadian DI benefits are optimal at conventional significance levels. Together, our empirical estimates indicate that DI benefits are valuable despite the relatively high costs of providing them. Based on the discussion in Section 2, where we argue that our approach likely provides a lower bound on the DI program's insurance value, we interpret our results as strong evidence that Canadian DI benefits are not too generous.

U.S. Implementation. Milligan and Schirle (2019) estimate regressions of the form $DI_{it} = \alpha + \beta_1 \cdot \ln(B_{it}) + \beta_2 \cdot \ln(E_{it}) + \epsilon_{it}$ to identify the effect of benefit generosity, B_{it} , and earnings, E_{it} , on DI take-up, DI_{it} . They instrument for B_{it} using a simulated benefits approach and employ a Bartik shift-share design for E_{it} . We apply our sufficient

statistics approach to their estimates to derive the implied insurance value and incentive costs of the U.S. DI program.

The DI take-up elasticity is $\varepsilon \equiv \frac{dDI/DI}{dB/B} = \frac{\beta_1}{DI}$. The estimates from Table 3, Column (4) in Milligan and Schirle (2019) for β_1 and the average DI rate from Table 1, Column (5) imply an elasticity of

$$\varepsilon = \frac{0.042}{0.037} = 1.14. \quad (16)$$

The fiscal multiplier is therefore $1 + \frac{\varepsilon}{1-\gamma} = 1 + \frac{1.14}{1-0.037} = 2.18$, i.e., a \$1 increase in U.S. DI benefits costs \$2.18.

To identify the insurance value with our approach we need $-\frac{\partial DI/\partial B}{\partial DI/\partial E}$. Table 3, Column (4) in Milligan and Schirle (2019) identifies the effect of a 100%-change in DI benefits and wages. We thus rescale the point estimates by the average DI benefit and average earnings from Table 1, Column (5): $\frac{\partial DI}{\partial B} = \frac{\beta_1}{B} = \frac{0.042}{15596}$ and $\frac{\partial DI}{\partial E} = \frac{\beta_2}{E} = \frac{-0.038}{48594}$. The implied insurance value is therefore

$$\text{Insurance Value} = -\frac{\partial DI/\partial B}{\partial DI/\partial E} = 3.44, \quad (17)$$

i.e., the additional dollar in DI benefits is valued at \$3.44, which exceeds the \$2.18 costs of providing the additional dollar.

7 Conclusion

In this paper we develop and implement a new approach to estimate the value of disability insurance from take-up decisions. We show that the take-up response to a change in disability benefits relative to a change in wages identifies the implied value of disability benefits in the classic DI model from Diamond and Sheshinski (1995). In models with richer heterogeneity, our method identifies the insurance value of the marginal applicants. We argue that in this case, our method identifies a lower bound for the average insurance value of DI benefits in the population. A key advantage of our approach, compared to existing methods for estimating the insurance value of social insurance programs, is that it can be implemented in a variety of contexts where data

on take-up decisions are available.

In our Canadian implementation, we find significant DI take-up effects for both increased DI benefit levels and negative economic shocks. However, the responses to benefit changes are quantitatively larger. Combining the wage and DI benefit estimates, we find that the insurance value of a \$1 increase in DI benefits is \$2.2 in our shift-share specification. We find even higher insurance values of \$2.8 to \$3.4 for the oil price shocks. For the incentive costs of higher DI benefits, we find that an extra dollar in disability benefits costs \$1.6 in total. Our estimates imply that the insurance value exceeds the fiscal costs in the Canadian DI program. Therefore, Canadian DI benefits are not too generous. Similarly, U.S. DI benefits are not too generous based on an application of our approach to existing U.S. estimates from Milligan and Schirle (2019).

References

- Abadie, A. (2003). Semiparametric Instrumental Variable Estimation of Treatment Response Models. *Journal of Econometrics* 113(2), 231–263.
- Autor, D. H. and M. G. Duggan (2003). The Rise in the Disability Rolls and the Decline in Unemployment. *Quarterly Journal of Economics* 118(1), 157–206.
- Baker, M. and D. Benjamin (1999). Early Retirement Provisions and the Labor Force Behavior of Older Men: Evidence from Canada. *Journal of Labor Economics* 17(4), 724–756.
- Baker, M. and K. Milligan (2012). Disability insurance programs in Canada. In D. A. Wise (Ed.), *Social security programs and retirement around the world: Historical trends in mortality and health, employment, and disability insurance participation and reforms*, pp. 327–58. Chicago:University of Chicago Press.
- Black, D., K. Daniel, and S. Sanders (2002). The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust. *American Economic Review* 92(1), 27–50.
- Bound, J. and R. V. Burkhauser (1999). Economic Analysis of Transfer Programs Targeted on People with Disabilities. In *Handbook of Labor Economics*, Volume 3, Chapter 51, pp. 3417–3528. Elsevier.
- Cabral, M. and M. R. Cullen (2019, August). Estimating the value of public insurance using complementary private insurance. *American Economic Journal: Economic Policy* 11(3), 88–129.
- Campolieti, M. and C. Riddell (2012). Disability Policy and the Labor Market: Evidence from a Natural Experiment in Canada, 1998–2006. *Journal of Public Economics* 96(3-4), 306–16.
- Charles, K. K., Y. Li, and S. J. Melvin (2018). Disability Benefit Take-Up and Local Labor Market Conditions. *The Review of Economics and Statistics*.
- Chetty, R. (2006). A General Formula for the Optimal Level of Social Insurance. *Journal of Public Economics* 90(10-11), 1879–1901.
- Chetty, R. (2008). Moral Hazard versus Liquidity and Optimal Unemployment Insurance. *Journal of Political Economy* 116(2), 173–234.

- Chetty, R. and A. Szeidl (2007). Consumption Commitments and Risk Preferences. *Quarterly Journal of Economics* 122(May), 831–877.
- Clausen, A. and C. Strub (2020). Reverse Calculus and Nested Optimization. *Journal of Economic Theory* 187, 105019.
- Deshpande, M. and L. M. Lockwood (2022). Beyond health: Nonhealth risk and the value of disability insurance. *Econometrica* 90(4), 1781–1810.
- Diamond, P. A. and E. Sheshinski (1995). Economic Aspects of Optimal Disability Benefits. *Journal of Public Economics* 57(1), 1–23.
- Diarra, M., G. Pelletier, and M. Pépin (2015). Évolution de la clientèle de la rente d’invalidité de 1970 à 2010. Technical report, Régie des Rentes du Québec.
- Einav, L. and A. Finkelstein (2011, March). Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic Perspectives* 25(1), 115–38.
- Einav, L., A. Finkelstein, and M. R. Cullen (2010). Estimating Welfare in Insurance Markets Using Variation in Prices. *Quarterly Journal of Economics* 123(August), 877–921.
- Fadlon, I. and T. H. Nielsen (2019). Household Labor Supply and the Gains from Social Insurance. *Journal of Public Economics* 171, 18–28.
- Frandsen, B., L. Lefgren, and E. Leslie (2023). Judging Judge Fixed Effects. *American Economics Review* 113(1), 253–77.
- Gruber, J. (1997). The Consumption Smoothing Benefits of Unemployment Insurance. *American Economic Review* 87(1), 192–205.
- Gruber, J. (2000). Disability Insurance Benefits and Labor Supply. *Journal of Political Economy* 108(6), 1162–1183.
- Haller, A., S. Staubli, and J. Zweimüller (2024). Designing disability insurance reforms: Tightening eligibility rules or reducing benefits? *Econometrica* 92(1), 79–110.
- Hendren, N. (2017, July). Knowledge of future job loss and implications for unemployment insurance. *American Economic Review* 107(7), 1778–1823.
- Landais, C. (2015). Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design. *American Economic Journal: Economic Policy* 7(4), 243–78.

- Landais, C. and J. Spinnewijn (2021). The Value of Unemployment Insurance. *The Review of Economic Studies* 88(6), 3041–3085.
- Low, H. and L. Pistaferri (2015). Disability Insurance and the Dynamics of the Incentive Insurance Trade-Off. *American Economic Review* 105(10), 2986–3029.
- Meyer, B. D. and W. K. Mok (2019). Disability, Earnings, Income and Consumption. *Journal of Public Economics* 171, 51–69.
- Milligan, K. and T. Schirle (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.
- Montiel Olea, J. and C. Pflueger (2013). A Robust Test for Weak Instruments. *Journal of Business and Economic Statistics* 31(3), 358–369.
- Seibold, A., S. Seitz, and S. Siegloch (2022). Privatizing disability insurance.
- Staubli, S. and Q. Zhao (2023). The Long-Run and Distributional Impacts of Lowering the Pension Eligibility Age. *Working Paper*.
- Stock, J. H. and M. Yogo (2005). Testing for Weak Instruments In Linear IV Regression. In D. W. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pp. 80–108. Cambridge: Cambridge University Press.
- Torjman, S. (2002). The Canada Pension Plan Disability Benefit. *Caledon Institute of Social Policy*.

APPENDIX

A Theoretical Framework: Extensions and Robustness

This Appendix discusses three extensions of the simple model from Section 2. First, Appendix A.1 introduces non-separability in the utility function and heterogeneity in wages, DI benefits, and unearned income. We show that in this case our approach identifies the insurance value of marginal applicants. Appendix A.1 then also discusses how the insurance value of the marginal applicant relates to the insurance value of the average DI recipient. Second, Appendix A.2 provides a model of endogenous job search with three labor market states: work, DI benefits, and other benefits. We show that we identify a lower bound on the insurance value if we identify the denominator in equation (11) from exogenous labor market shocks that shift both wages and employment probabilities. Third, Appendix A.3 provides a dynamic model to address the question how we can make permanent DI benefit changes comparable to temporary wage changes.

A.1 Non-Separability and Heterogeneity

The two critical assumptions that the relative take-up responses identify the insurance value of DI benefits are that (i) the disability level enters the utility function additively and (ii) individuals only differ in their disability level. The two assumptions ensure that the insurance values of the marginal applicant and the average DI recipient are the same. We relax both assumptions here, showing that with non-separability or richer heterogeneity, our approach identifies the insurance value for marginal applicants.

Non-Separability of Consumption and Disutility of Work. The classic DI model in Diamond and Sheshinski (1995) assumes that disability (or disutility of work) enters the utility function additively: utility while working is $u(c^w) - \theta$ in Diamond and Sheshinski (1995). More general utility functions would be of the form $u(c^w, \theta)$ and $v(c^b, \theta)$. The application decision is then: apply if $v(c^b, \theta) - u(c^w, \theta) \geq 0$. As long as there is

monotonicity, i.e., more disabled individuals are more likely to apply, there is again a marginal applicant determined by the solution to $H = v(c^b, \theta^A) - u(c^w, \theta^A) = 0$. We can then apply the implicit function theorem, and the relative DI take-up effects are given by

$$-\frac{\partial DI}{\partial b} / \frac{\partial DI}{\partial w} = \frac{v'(c^b, \theta^A)}{u'(c^w, \theta^A)}. \quad (\text{A.1})$$

Thus, our approach still measures the insurance value of marginal applicants, but it may no longer be representative for the average DI recipient's insurance value.

Heterogeneity. To introduce heterogeneity beyond disability levels, we let after-tax wages $w_i - \tau_i$, DI benefits b_i , unearned income A_i , and disability level θ_i vary across individuals i . Consumption of individual i when working is then given by $c_i^w = w_i - \tau_i + A_i$ and consumption when on DI benefits is $c_i^b = b_i + A_i$. Assume there is a distribution of wages, DI benefits, and unearned income $G(w, b, A)$ and a conditional distribution of disability $F(\theta|w, b, A)$. The marginal applicant for a given wage, DI benefit, and unearned income is given by

$$\theta^A(w, b, A) = u(c^w) - v(c^b). \quad (\text{A.2})$$

DI take-up is determined by $DI = \int \int_{\theta^A(w, b, A)}^{\infty} p(\theta) dF(\theta|w, b, A) dG(w, b, A)$. Therefore,

$$-\frac{dDI}{db} / \frac{dDI}{dw} = \frac{\int [v'(c^b) f(\theta^A) p(\theta^A) db] dG(w, b, A)}{\int [u'(c^w) f(\theta^A) p(\theta^A) dw] dG(w, b, A)}. \quad (\text{A.3})$$

The relative DI take-up responses with respect to DI benefits and wages still identify the insurance value of marginal applicants—the ratio of the marginal utilities of consumption on DI benefits and while working—but it could again differ from the insurance value of the average DI recipient.

Composition or Insurance Value? When implementing equation (A.3), an empirical challenge arises. The empirical variation we use to identify DI take-up effects of benefits and wages might not affect the same individuals. In this case, the differential take-up

responses we identify empirically could be driven by composition (different people are affected) rather than by differences in marginal utilities between working and receiving DI benefits. We address this concern in Section 5 by implementing a complier analysis that compares the characteristics of marginal applicants for changes in benefits with those of marginal applicants for changes in wages. Reassuringly, Table 3 shows that the incidence of these two groups of marginal applicants across the earnings and taxable income distribution is strikingly similar. Moreover, mortality, a measure of disability severity, is not statistically different between the two complier groups. These findings indicate that the empirical variation in DI benefits and wages that we exploit affect similar (w, A, θ) -types, suggesting we identify meaningful differences in marginal utilities and not composition effects.¹³

Is the Insurance Value of the Marginal Applicant a Lower Bound? A natural question is how representative is the insurance value of the marginal applicant for the insurance value of the average DI recipient? Intuitively, one would expect the insurance value of the marginal applicant to be a lower bound for the insurance value of all DI applicants. But it is not obvious whether this intuition is accurate. What we want to measure is the value of a transfer between DI and non-DI recipients, which we can write as

$$IV_{pop} = \frac{E [v'(c^b(\theta); \theta)|\text{on DI}]}{E [u'(c^w(\theta); \theta)|\text{not on DI}]}, \quad (\text{A.4})$$

where we allow consumption and utility itself to vary by the disability level. Our approach identifies the value of a transfer between the DI and the non-DI state for marginal applicants

$$IV_A = \frac{E [v'(c^b(\theta); \theta)|\theta^A]}{E [u'(c^w(\theta); \theta)|\theta^A]}. \quad (\text{A.5})$$

The question is whether $IV_A \leq IV_{pop}$. It is reasonable to expect that marginal applicants have higher work capacity and thus potentially higher consumption than the average DI recipient, $c^b(\theta^A) \geq E [c^b(\theta)|\text{on DI}]$. It is also reasonable to expect that marginal applicants have lower work capacity than the average non-DI recipient,

¹³Since DI benefits are a deterministic function of past earnings, we have comparable (w, A, b, θ) -types.

$c^w(\theta^A) \leq E [c^w(\theta)|\text{not on DI}]$. If the marginal utility of consumption does not depend on θ and the utility functions are concave, we have

$$\frac{E [v'(c^b(\theta))|\theta^A]}{E [u'(c^w(\theta))|\theta^A]} \leq \frac{E [v'(c^b(\theta))|\text{on DI}]}{E [u'(c^w(\theta))|\text{not on DI}]} \quad (\text{A.6})$$

and the insurance value of marginal applicants indeed provides a lower bound for the true insurance value of DI benefits.

If marginal utility of consumption varies directly with the disability level, we need to consider two scenarios. The first scenario is that marginal utility of consumption is higher for the more disabled, which is the modeling assumption in the structural DI model of Low and Pistaferri (2015). In this case, $v'(c^b; \theta^A) \leq E [v'(c^b; \theta)|\text{on DI}]$ and $u'(c^w; \theta^A) \geq E [u'(c^w; \theta)|\text{not on DI}]$ as long as more disabled individuals are more likely to end up on DI benefits, i.e., if $(E [\theta|\text{not on DI}] \leq \theta^A \leq E [\theta|\text{on DI}])$ holds. Hence, the marginal applicants' insurance value is a lower bound in this scenario.

The second scenario is that marginal utility of consumption is lower for more disabled. Then, the insurance value of the marginal applicants might not be a lower bound for the actual insurance value. It is not obvious if and how disability severity affects marginal utility of consumption, but falling marginal utility in disability severity implies that DI benefits should fall with disability severity. No DI program features such a benefit schedule. In DI programs where the benefit level varies with disability severity, it is increasing in severity.

Based on these arguments, it is reasonable that the insurance value we estimate for marginal applicants provides a lower bound for the actual insurance value of DI benefits.

A.2 Model with Three Labor Market States

We now extend the model with a third labor market state—other benefits—and endogenous job search to address an empirical challenge. The third labor market state does not change our results from a theoretical perspective. The relative DI responses to changes in benefits and wages still identify the insurance value. However, empirically, it is challenging to isolate pure wage effects. The shift-share instruments we use affect

wages but can also influence separation and job-finding rates.

In the extended model, individuals are employed with probability $s(e; \Omega)$, which depends on individual effort e and exogenous economic conditions Ω . Individuals are unemployed with probability $1 - s(e; \Omega)$. An individual's wage, $w(e; \Omega)$, is also a function of individual effort and exogenous economic conditions. Effort is costly and creates utility loss $\psi(e; \theta)$, which varies by the disability level θ . The optimal effort level, e , is determined by

$$\max_e s(e; \Omega) \cdot u(w(e; \Omega)) + (1 - s(e; \Omega)) \cdot v(z) - \psi(e; \theta) \quad (\text{A.7})$$

An interior solution fulfills the following first-order condition

$$s'(e; \Omega) \cdot [u(w(e; \Omega)) - v(z)] + s(e; \Omega) \cdot u'(w; \Omega) \cdot \frac{\partial w(e; \Omega)}{\partial e} - \psi'(e; \theta) = 0. \quad (\text{A.8})$$

Individuals with too high θ might choose a corner solution with $e = 0$. In this case, we assume that $s(0) = 0$ and individuals receive other welfare benefits z with certainty.

If individuals are awarded DI, they leave the labor force permanently and do not incur any effort cost/disutility; they have flow utility $v(b)$. Absent application costs, an individual applies for DI benefits if the expected utility of receiving DI exceeds the expected utility of being in the labor force:

$$v(b) \geq s(e; \Omega) \cdot u(w) + (1 - s(e; \Omega)) \cdot v(z) - \psi(e; \theta) \quad (\text{A.9})$$

The marginal applicant θ^A is determined by

$$\Theta \equiv s(e; \Omega) \cdot u(w) + (1 - s(e; \Omega)) \cdot v(z) - \psi(e; \theta^A) - v(b) = 0. \quad (\text{A.10})$$

From (A.10), a change in DI benefits has the following impact on application behavior:

$$\frac{\partial \theta^A}{\partial b} = \frac{v'(b)}{\partial \Theta / \partial \theta^A}. \quad (\text{A.11})$$

An economic shock Ω has the following impact on application behavior:

$$\frac{\partial \theta^A}{\partial \Omega} = -\frac{1}{\partial \Theta / \partial \theta^A} \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [u(w) - v(z)] + s \cdot u'(w) \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]. \quad (\text{A.12})$$

A negative economic shock thus increases applications through two channels: (i) increased job loss $\frac{\partial s(e; \Omega)}{\partial \Omega}$ and (ii) reduced wages $\frac{\partial w(e; \Omega)}{\partial \Omega}$. The ratio of the take-up response to changes in benefits and economic conditions is given by

$$\frac{\frac{\partial \theta^A}{\partial b}}{-\frac{\partial \theta^A}{\partial \Omega}} = \frac{v'(b)}{\left[\frac{\partial s(e; \Omega)}{\partial \Omega} [u(w) - v(z)] + s \cdot u'(w) \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]} \quad (\text{A.13})$$

$$\leq \frac{v'(b)}{u'(w) \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [w - z] + s \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]}. \quad (\text{A.14})$$

The inequality in (A.14) holds if $u'(w)(w - z) \leq u(w) - v(z)$. That is, the monetized utility loss associated with job loss, $(u(w) - v(z))/u'(w)$, is at least as large as the income loss associated with job loss, $(w - z)$. If the utility function is not state-dependent, i.e., $u(\cdot) = v(\cdot)$, and the replacement rate of other benefits is less than 100 percent, $w \geq z$, the condition holds for concave utility functions (falling marginal utility of consumption). If we assume $u'(w)(w - z) \leq u(w) - v(z)$, we can rescale the economic shock, $\partial \Omega$, by its impact on income (=earnings plus benefits) $\frac{\partial}{\partial \Omega} [s \cdot w + (1 - s) \cdot z] = \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [w - z] + s \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]$ and get

$$\frac{\frac{\partial \theta^A}{\partial b}}{-\frac{\partial \theta^A}{\partial \Omega} / \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [w - z] + s \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]} \leq \frac{v'(b; \theta^A)}{u'(w; \theta^A)} \leq \frac{E[v'(b; \theta)|\text{on DI}]}{E[u'(w; \theta)|\text{working}]} \quad (\text{A.15})$$

Hence, our empirical estimate that rescales the inflow effect by income is a lower bound for the insurance value of the marginal applicant and a lower bound for the program's insurance value.

A.3 Dynamic Model

We now extend the model to a dynamic setting. We first discuss the dynamic version of the model from the main text with two labor market states in Section A.3.1. We then discuss the dynamic version of the model from Appendix A.2 with three labor market states in Section A.3.2. The findings from the static model generalize to the dynamic setting. The challenge in the dynamic setup is empirical because we need to make temporary changes in wages comparable to permanent changes in DI benefits.

A.3.1 Dynamic Model with Two Labor Market States.

Consider the dynamic version of the simple static model from Section 2. We extend the model to T periods and allow for richer heterogeneity beyond θ . θ and other state variables evolve stochastically over the agent's relevant time horizon. Let $X_{i,t} = \{\theta_{i,t}, A_{i,t}, \chi_{i,t}\}$ denote the vector of state variables where $\theta_{i,t}$ denotes agent i 's disability level in period t , $A_{i,t}$ denotes the asset level, and $\chi_{i,t}$ is a vector of other state variables (allowing for heterogeneity across agents in other dimensions like differences in education or experience that translate to differences in wages). The state vector $X_{i,t}$ summarizes all the information relevant for agent i 's choices in period t . The laws of motion of assets in the disability and employment state are

$$A_{i,t+1} = (1 + r_t)A_{i,t} + b(X_{i,t}) - c_{i,t}^b(X_{i,t}) \quad (\text{A.16})$$

$$A_{i,t+1} = (1 + r_t)A_{i,t} + w(X_{i,t}) - \tau_{i,t}(X_{i,t}) - c_{i,t}^w(X_{i,t}), \quad (\text{A.17})$$

where DI benefits of individual i in period t , $b(X_{i,t})$, can depend on the agent's state vector $X_{i,t}$. Analogously, $w(X_{i,t})$ denotes labor income and $\tau(X_{i,t})$ are taxes. Agents make state contingent plans on how much to consume in each labor market state $\{c_{i,t}^b(X_{i,t}), c_{i,t}^w(X_{i,t})\}$, and whether they apply to DI benefits $\alpha_{i,t}(X_{i,t}) \in \{0, 1\}$.

The sequence of events and choices is as follows. At the beginning of the period, the shocks $\theta_{i,t}$ and $\chi_{i,t}$ are revealed. Having learned $X_{i,t}$, the individual decides whether to apply to DI, $\alpha_{i,t}(X_{i,t}) = 1$, or not, $\alpha_{i,t}(X_{i,t}) = 0$. If the application is accepted, individuals become DI beneficiaries in the next period, $t + 1$, for the rest of their life. If the application is rejected, individuals stay in the labor market and can apply again the

next period.

Denote by $DI_{i,t+1}$ the probability that an agent i is a DI benefit recipient in period $t + 1$:

$$DI_{i,t+1} = 1 - \left[\prod_{k=0}^t (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k})) \right]. \quad (\text{A.18})$$

The probability that agent i transitions to DI in period $k + 1$ is $\alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k})$. The other state variables, disutility of work $\theta_{i,t}$ and $\chi_{i,t}$, follow stochastic processes that can, in principle, depend on agents' choices. The expectation operator $E[\cdot]$ below captures the evolution of the state variables.¹⁴ The agent's problem is given by

$$\begin{aligned} V_i(b) = \max E & \left[\sum_{t=0}^{T-1} \beta^t \left(v(c_{i,t}^b) \cdot D_{i,t} + \left(u(c_{i,t}^w) - \psi(\theta_{i,t}) \right) \cdot (1 - D_{i,t}) \right) \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \left((1 + r_t) A_{i,t} + b_{i,t} - c_{i,t}^b - A_{i,t+1} \right) D_{i,t} \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^W \left((1 + r_t) A_{i,t} + w_{i,t} - \tau_{i,t} - c_{i,t}^w - A_{i,t+1} \right) (1 - D_{i,t}) \right] \end{aligned} \quad (\text{A.19})$$

Planner's Problem. The social planner maximizes social welfare by choosing DI benefits and solves

$$\max_b W(b) = \int_i V_i(b) di + \lambda \cdot G(b) \quad (\text{A.20})$$

where

$$G(b) = \int_i E \left[\sum_{t=0}^{T-1} (1 + r_t)^{-t} \left((1 - D_{i,t}) \cdot \tau_{i,t} - D_{i,t} \cdot b_{i,t} \right) \right] di \quad (\text{A.21})$$

is the planners net revenue and λ denotes the Lagrange multiplier on the planner's budget constraint.

¹⁴The operator $E[Y]$ aggregates the variable Y over states of nature, i.e., $E[Y] = \int Y(X_{i,t}) dF(X_{i,t})$ where $F(\cdot)$ is the distribution of the state variables $X(i, t)$. This formulation is flexible: the only restriction we impose on the distribution of state variables is that it does not directly depend on DI benefits $\{b_t\}_{t=0}^{T-1}$ itself. The evolution of $X_{i,t}$, however, can depend on the agent's choices, which themselves depend on DI benefits.

Optimal DI Benefit. We focus on permanent and uniform DI benefits: $b(X_{i,t}) = b$ for all $X_{i,t}$. This resembles our empirical setup where the flat DI benefit uniformly changed for all DI recipients.¹⁵ The optimal DI benefit solves

$$\frac{\partial W(b)}{\partial b} = E \left[\int_i \sum_{t=0}^{T-1} \beta^t \left(v'(c_{i,t}^b) \cdot D_{i,t} \right) di \right] + \lambda \cdot \frac{\partial G(b)}{\partial b} = 0 \quad (\text{A.22})$$

where the budget effect is

$$\frac{\partial G(b)}{\partial b} = - \int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} \left(D_{i,t} + \frac{\partial D_{i,t}}{\partial b} \cdot (b+\tau) \right) \right] di \quad (\text{A.23})$$

$$= -DI \left[1 + \frac{\partial DI}{\partial b} (b+\tau) \right] \quad (\text{A.24})$$

$$= -DI \left[1 + \frac{\partial DI}{\partial b} \frac{b}{DI} \frac{1}{1-DI} \right] \quad (\text{A.25})$$

and $DI \equiv \int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} (D_{i,t}) \right] di$ is the average DI recipient share in the population.¹⁶ We assume that the planner's budget constraint is differentiable, i.e., $\frac{\partial G(b)}{\partial b}$ exists.

Proof. of (A.22).

The first order condition in (A.22) follows from applying Clausen and Strub (2020)'s Differentiable Sandwich Lemma. Clausen and Strub (2020) show that if a function $F(b)$ is sandwiched at some point b between two differentiable functions (upper and lower support functions $U(b)$ and $L(b)$), then this function F is differentiable at this point b and the derivative of the sandwiched function F equals the derivative of the upper and lower support functions at this point, i.e., $F'(b) = U'(b) = L'(b)$. Suppose b^* is the DI benefit level that maximizes welfare. By definition $W(b^*) \geq W(b) \forall b$. We take the constant function $U(b) = W(b^*)$ as the upper support function with $U'(b) = 0$.

¹⁵The more general problem, where the planner chooses the optimal benefit function $b_s(X_{i,s})$ in each period s and each $X_{i,s}$, poses no additional theoretical challenges. Similar formulas apply. The only difference is that the incidence of the benefit changes is then no longer uniform across all DI recipients. For notational simplicity, we abstract from time-varying benefits since it is irrelevant in our empirical implementation with a uniform DI benefit change.

¹⁶Equation (A.25) uses the balanced budget constraint, implying $\tau = \frac{DI}{1-DI} \cdot b$.

For the lower support function, we take $L(b) = \int \bar{V}_i(b) di + \lambda (G(b))$ where $\bar{V}_i(b)$ denotes the agent's indirect utility if she sticks to her behavior that is optimal for benefits b^* even when benefit levels change. Thus, $\bar{V}_i(b) \leq V_i(b)$ and $L(b)$ is a lower support function for $W(b)$, which implies that

$$\frac{\partial W(b)}{\partial b} = \frac{\partial L(b)}{\partial b} \quad (\text{A.26})$$

$$= E \left[\int_i \sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \cdot D_{i,t} di \right] + \lambda \cdot \frac{\partial G(b)}{\partial b} \quad (\text{A.27})$$

$$= E \left[\int_i \sum_{t=0}^{T-1} \beta^t \left(v'(c_{i,t}^b) \cdot D_{i,t} \right) di \right] + \lambda \cdot \frac{\partial G(b)}{\partial b} \quad (\text{A.28})$$

$$= \frac{\partial U(b)}{\partial b} = 0 \quad (\text{A.29})$$

and concludes the proof of (A.22). \square

If we assume again that $\lambda = \mathbb{E} [u'(c^w) | DI = 0] \equiv \int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} (1-D_{i,t}) \cdot u'(c_{i,t}^w) \right] di / \left[\int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} (1-D_{i,t}) \right] di \right]$, we arrive from (A.22) at the Baily-Chetty formula

$$\frac{\mathbb{E} [v'(c^b) | DI = 1]}{\mathbb{E} [u'(c^w) | DI = 0]} = 1 + \frac{\varepsilon_{DI,b}}{1-DI}, \quad (\text{A.30})$$

where $\mathbb{E} [v'(c^b) | DI = 1] \equiv \int_i E \left[\sum_{t=0}^{T-1} \beta^t \cdot v'(c_{i,t}^b) \cdot D_{i,t} \right] di / \left[\int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} (D_{i,t}) \right] \right]$ denotes the average marginal utility of consumption of DI recipients.

Application Decision. In the dynamic setting, the decision to apply for DI benefits becomes forward-looking. At time t the value function of being employed is

$$V_{i,t}^E (X_{i,t}) = u(c_{i,t}^w) + \beta \cdot E \left[\alpha_{i,t} p_{i,t} V_{i,t+1}^D (X_{i,t+1}) + (1 - \alpha_{i,t} p_{i,t}) V_{i,t+1}^E (X_{i,t+1}) | X_{i,t} \right] - \psi(\theta_{i,t}). \quad (\text{A.31})$$

Individuals decide in period t whether to apply for DI benefits, $\alpha_{i,t} = 1$, or not apply, $\alpha_{i,t} = 0$. An applicant is awarded benefits with probability $p_{i,t}$. In case of an award, individuals receive DI benefits in the next period. Hence, individuals employed in period t enter the DI system with probability $\alpha_{i,t} p_{i,t}$ in period $t+1$ and remain employed with

probability $1 - \alpha_{i,t}p_{i,t}$. The value of being on DI, which is an absorbing state, is given by

$$V_{i,t}^D(X_{i,t}) = v(c_{i,t}^b) + \beta \cdot E[V_{i,t+1}^D | X_{i,t}]. \quad (\text{A.32})$$

In period t individuals apply for DI benefits, $\alpha_{i,t} = 1$, if the expected value of receiving DI benefits exceeds the expected value of working:

$$\Gamma(\theta_{i,t}, A_{i,t}, \chi_{i,t}) \equiv E[V_{i,t+1}^D - V_{i,t+1}^E | \theta_{i,t}, A_{i,t}, \chi_{i,t}] \geq 0. \quad (\text{A.33})$$

For given state variables A_t and χ_t , the marginal applicant θ_t^A in period t is defined by

$$\Gamma(\theta_t^A, A_t, \chi_t) = 0. \quad (\text{A.34})$$

$\Gamma(\theta_t^A, A_t, \chi_t) = 0$ defines a unique marginal applicant for given A_t and χ_t if the disability shock is persistent. By persistence we mean that a disability shock today reduces the expected value of working, i.e., $\frac{\partial E[V_{i,t+1}^E | X_{i,t}]}{\partial \theta_t} < 0$.¹⁷ In this case, θ_t^A conditional on A_t and χ_t is uniquely defined by (A.34), because $E[V_{i,t+1}^D | X_{i,t}]$ does not depend on the disability level itself, i.e., $\frac{\partial E[V_{i,t+1}^D | X_{i,t}]}{\partial \theta_t} = 0$, while $\frac{\partial E[V_{i,t+1}^E | X_{i,t}]}{\partial \theta_t} < 0$ and hence $E[V_{i,t+1}^D | X_{i,t}]$ and $E[V_{i,t+1}^E | X_{i,t}]$ have single crossing.

Inflow Responses Reveal Insurance Value of DI. Equation (A.34) implicitly defines the marginal applicant at age t conditional on the state variables A_t and χ_t . We assume $\frac{\partial E[V_{i,t+1}^E | X_{i,t}]}{\partial \theta_t} < 0$, i.e., the value of continuing working varies smoothly with the disability level ($E[V_{i,t+1}^E | X_{i,t}]$ is differentiable in θ_t) and a disability shock today reduces

¹⁷Persistence in disability shocks makes intuitive sense. The alternatives to persistence are harder to motivate. If a disability shock today does not change the expected value of working, $\frac{\partial E[V_{i,t+1}^E]}{\partial \theta_t} = 0$, then everybody would always apply, or everybody would never apply (depending on whether working is more attractive than receiving DI benefits, on average). The application decision would thus be independent of the disability level. If a disability shock increases the value of working $\frac{\partial E[V_{i,t+1}^E]}{\partial \theta_t} > 0$, then the application decision would be positively correlated with health, i.e., positive health shocks today would increase DI applications and negative health shocks today would decrease the application probability.

the expected value of working, implying that θ_t^A is differentiable in b by the implicit function theorem. A change in the DI benefit b has the following DI inflow effect at age $t + 1$:

$$\frac{\partial}{\partial b} DI_{t+1} = \frac{\partial}{\partial b} \int E \left[\int_{\theta_t^A}^{\infty} p(\theta) dF(\theta) | A_t, \chi_t \right] dG(A_t, \chi_t) \quad (\text{A.35})$$

$$= \int E \left[-\frac{\partial \theta_t^A}{\partial b} p(\theta_t^A) f(\theta_t^A) | A_t, \chi_t \right] dG(A_t, \chi_t), \quad (\text{A.36})$$

where $G(A_t, \chi_t)$ denotes the distribution of the state variables A_t and χ_t . Note that there can be multiple marginal applicants so that different disability cutoffs θ_t^A apply for different state variables A_t and χ_t . Equation (A.36) sums up the responses from all marginal applicants. Furthermore, we take the distribution $G(A_t, \chi_t)$ as given, which means that equation (A.36) quantifies the DI inflow effect at age $t + 1$ when individuals learn at age t that DI benefits permanently change. This assumption reflects our empirical approach that sums up inflow effects at different ages from different cohorts in the short run.¹⁸ Equation (A.34), the implicit function theorem, and the envelope theorem imply that

$$\frac{\partial \theta_t^A}{\partial b} = - \frac{\frac{\partial E [V_{i,t+1}^D - V_{i,t+1}^E | \theta_t^A, A_t, \chi_t]}{\partial b}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.37})$$

$$= - \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) | \theta_t^A, A_t, \chi_t \right] \quad (\text{A.38})$$

¹⁸Empirically we cannot estimate the full life-cycle effect of a permanent DI benefit change within one generation, which corresponds to the thought experiment in equation (A.22). However, we argue below that our empirical estimates still identify a lower bound for the insurance value.

where $1 - DI_{f,s} \equiv \prod_{j=f}^s (1 - \alpha_j p_j)$. Combining (A.36) and (A.38) yields

$$\begin{aligned} \frac{\partial}{\partial b} DI_{t+1} &= \int E \left[c(\theta_t^A) \left(v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in Q^A \right], \end{aligned} \quad (\text{A.39})$$

where $c(\theta_t^A) \equiv \frac{p(\theta_t^A) f(\theta_t^A)}{\partial \Gamma_t / \partial \theta_t^A}$, and $\delta_{t+1}^k \equiv \beta^k (1 - DI_{t+1,t+k})$. Therefore, the inflow effect at age $t + 1$ corresponds to the ($c(\theta_t^A)$ -weighted) average marginal utility of consumption of all marginal applicants at age t when on DI benefits, which we denote by expression (A.39).

Analogously, a permanent change in wages implies

$$\frac{\partial \theta_t^A}{\partial w} = - \frac{\frac{\partial E [v_{i,t+1}^D - v_{i,t+1}^E | \theta_t^A, A_t, \chi_t]}{\partial w}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.40})$$

$$= \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) u'(c_{t+1+k}^w) | \theta_t^A, A_t, \chi_t \right] \quad (\text{A.41})$$

and the inflow effect at age $t + 1$ is given by

$$\begin{aligned} \frac{\partial}{\partial w} DI_{t+1} &= - \int E \left[c(\theta_t^A) \left(u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) u'(c_{t+1+k}^w) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv -E \left[u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k u'(c_{t+1+k}^w) | \theta_t^A \in Q^A \right]. \end{aligned} \quad (\text{A.42})$$

Our empirical estimate of the DI inflow effect sums up the inflow effect at all ages, i.e.,

$$\frac{\partial DI / \partial b}{-\partial DI / \partial w} = \frac{\sum_{t=0}^{T-2} \frac{\partial}{\partial b} DI_{t+1}}{-\sum_{t=0}^{T-2} \frac{\partial}{\partial w} DI_{t+1}} \quad (\text{A.43})$$

$$= \frac{\sum_{t=0}^{T-2} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in Q^A \right]}{\sum_{t=0}^{T-2} E \left[u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k u'(c_{t+1+k}^w) | \theta_t^A \in Q^A \right]} \quad (\text{A.44})$$

Equation A.44 indicates that the ratio of DI inflow responses with respect to changes

in benefits and wages corresponds to the marginal applicants' ratio of marginal utilities of consumption when on DI benefits versus when working. The ratio of relative inflow responses provides a lower bound for the insurance value of DI benefits if marginal applicants' insurance value is lower in all periods than that of all DI recipients (as we assume in the static model).¹⁹

Temporary Changes in Wages. We now consider the DI inflow response to temporary wage changes. For a temporary wage change $w_{T,t} = \{dw_{i,k}\}_{k=t}^{T-1}$, where the intensity of the shock $dw_{i,k}$ can vary over time and disappear after some time ($dw_{i,t>s} = 0$), we have

$$\begin{aligned} \frac{\partial \theta_t^A}{\partial w_{T,t+1}} &= - \frac{\frac{\partial E \left[v_{i,t+1}^D - v_{i,t+1}^E | \theta_t^A, A_t, \chi_t \right]}{\partial w_{T,t+1}}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \\ &\approx \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[u'(c_{t+1}^w) dw_{i,t+1} + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) u'(c_{t+1+k}^w) dw_{i,t+1+k} | \theta_t^A, A_t, \chi_t \right]. \end{aligned} \quad (\text{A.45})$$

To make the response of a temporary wage shocks comparable to a permanent benefit change, we rescale the response by the present value of the wage shock $PVW_{t+1} = dw_{i,t+1} + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) dw_{i,t+1+k}$. This formulation corresponds to our empirical approach, where we rescale the effects of local labor market shocks by their impact on lifetime income. The question is how this rescaled effect, $\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_{T,t}}$, compares to the effect of a permanent wage shock, $\frac{\partial \theta_t^A}{\partial w}$. Both $\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_{T,t}}$ and $\frac{\partial \theta_t^A}{\partial w}$ measure weighted average marginal utilities of consumption of the marginal applicants. The dif-

¹⁹We can use the mediant theorem to show this statement formally. Equation (A.44) is a sum of the form $\frac{\sum v_j}{\sum u_j}$, where $v_j \equiv c(\theta_{t(j)}^A) \delta_{t(j)+1}^k v'(c_{t(j)+1+k(j)}^b)$ and $u_j \equiv c(\theta_{t(j)}^A) \delta_{t(j)+1}^{k(j)} u'(c_{t(j)+1+k(j)}^w)$ for all t, k -combinations in (A.44) for all marginal applicants (including $k = 0$, where $v_j \equiv c(\theta_{t(j)}^A) v'(c_{t(j)+1}^b)$ and $u_j \equiv c(\theta_{t(j)}^A) u'(c_{t(j)+1}^w)$). Let $q_j = \frac{v_j}{u_j} = \frac{v'(c_{t(j)}^b)}{u'(c_{t(j)}^w)}$ and the indexing j is such that $q_1 \leq \dots \leq q_j \leq \dots \leq q_N$. By the mediant theorem, $\frac{\partial DI / \partial b}{\partial DI / \partial w} \leq q_N$, and so if the marginal applicants' insurance value is always lower than that of all DI recipients, we have $\frac{\partial DI / \partial b}{\partial DI / \partial w} \leq q_N \leq \frac{\mathbb{E}[v'(c^b) | DI=1]}{\mathbb{E}[u'(c^w) | DI=0]}$.

ference is that $\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T}$ puts more weight on periods in the nearer future when the temporary wage shocks are present.

We have $-\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T} \approx -\frac{\partial \theta_t^A}{\partial w}$ if marginal utility of consumption while working is roughly constant across periods for the marginal applicants. When marginal utility of consumption is falling with age, we have $-\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T} \geq -\frac{\partial \theta_t^A}{\partial w}$. If $-\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T} \geq -\frac{\partial \theta_t^A}{\partial w}$ holds, then the rescaled inflow response to a temporary wage shock is larger than the inflow response to a permanent wage shock. Following the steps above, it then follows that $\frac{\partial DI/\partial b}{\partial DI/\partial w_T} \leq \frac{\partial DI/\partial b}{\partial DI/\partial w} \leq \text{Insurance Value}$ and our empirical implementation provides a lower bound for the insurance value of DI benefits. If $-\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T} < -\frac{\partial \theta_t^A}{\partial w}$ holds, then $\frac{\partial DI/\partial b}{\partial DI/\partial w_T}$ is likely still a lower bound for the insurance value. As long as the weighted marginal utility of consumption of marginal applicants that enters into $\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T}$ is larger than the average marginal utility of consumption in the working population, we have $\frac{\partial DI/\partial b}{\partial DI/\partial w_T} \leq \frac{\mathbb{E}[v'(c^b)|DI=1]}{\mathbb{E}[u'(c^w)|DI=0]} = \text{Insurance Value}$.

A.3.2 Dynamic Model with Three Labor Market States.

Setup. We now extend the model with three labor market states from Appendix A.2 to a dynamic setting. The same notation, timing, and constraints from the dynamic model in Appendix A.3.1 apply. But there is an additional labor market state, unemployment, and decisions about job search and effort on the job have to be made. The laws of motion of assets in the disability, employment and unemployment state are

$$A_{i,t+1} = (1 + r_t)A_{i,t} + b(X_{i,t}) - c_{i,t}^b(X_{i,t}) \quad (\text{A.46})$$

$$A_{i,t+1} = (1 + r_t)A_{i,t} + w(X_{i,t}) - \tau_{i,t}(X_{i,t}) - c_{i,t}^w(X_{i,t}) \quad (\text{A.47})$$

$$A_{i,t+1} = (1 + r_t)A_{i,t} + z(X_{i,t}) - c_{i,t}^z(X_{i,t}) \quad (\text{A.48})$$

where $z(X_{i,t})$ denote unemployment benefits.

The sequence of events and choices is the following. At the beginning of the period, the shocks $\theta_{i,t}$ and $\chi_{i,t}$ are revealed. Having learned $X_{i,t}$, which also includes the labor market state (employed, unemployed, on disability benefits), the individual decides whether to apply to DI, $\alpha_{i,t}(X_{i,t}) = 1$, or not, $\alpha_{i,t}(X_{i,t}) = 0$. If the application is rejected, individuals stay in the labor market and can apply again the next period. Moreover, indi-

viduals also decide on how hard to search for a job if unemployed and how much effort to exert on the job. With effort $e_{i,t}$ unemployed individuals are employed next period with probability $q(e_{i,t}; \Omega_t)$ that depends both on their search effort and the exogenous economic conditions Ω_t . Employed individuals remain in their job in the next period with probability $s(e_{i,t}; \Omega_t)$, which again depends on on-the-job effort $e_{i,t}$ and economic conditions Ω_t .

Denote by $DI_{i,t+1}$ the probability that in period $t + 1$ agent i is a DI benefit recipient, which is

$$DI_{i,t+1} = 1 - \left[\prod_{k=0}^t (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k})) \right]. \quad (\text{A.49})$$

We denote by $E_{i,t+1}$ the probability that agent i is employed in period $t + 1$, which is

$$E_{i,t+1} = (1 - DI_{i,t+1}) (s(e_{i,t}; \Omega_t) + q(e_{i,t}; \Omega_t)). \quad (\text{A.50})$$

The probability to be unemployed is $U_{i,t+1} = 1 - DI_{i,t+1} - E_{i,t+1}$. The agent's problem is given by

$$\begin{aligned} V_i(b) = \max E & \left[\sum_{t=0}^{T-1} \beta^t \left(v(c_{i,t}^b) \cdot DI_{i,t} + \left(u(c_{i,t}^w) - \psi(\theta_{i,t}, e_{i,t}) \right) \cdot E_{i,t} + \left(u(c_{i,t}^z) - \phi(\theta_{i,t}, e_{i,t}) \right) \cdot U_{i,t} \right) \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \left((1 + r_t) A_{i,t} + b_{i,t} - c_{i,t}^b - A_{i,t+1} \right) DI_{i,t} \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^W \left((1 + r_t) A_{i,t} + w_{i,t} - \tau_{i,t} - c_{i,t}^w - A_{i,t+1} \right) E_{i,t} \right] \\ & + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^Z \left((1 + r_t) A_{i,t} + z_{i,t} - c_{i,t}^z - A_{i,t+1} \right) U_{i,t} \right]. \end{aligned} \quad (\text{A.51})$$

Planner's Problem. The social planner maximizes social welfare by choosing DI benefits and solves

$$\max_b W(b) = \int_i V_i(b) di + \lambda \cdot G(b) \quad (\text{A.52})$$

where

$$G(b) = \int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} (E_{i,t} \cdot \tau_{i,t} - D_{i,t} \cdot b_{i,t} - U_{i,t} \cdot z_{i,t}) \right] di \quad (\text{A.53})$$

is the planners net revenue and λ denotes the Lagrange multiplier on the planner's budget constraint. The solution to the planner's problem is still governed by (A.22). The only difference compared to the two labor market state model is that potential spillover effects to unemployment benefits enter the fiscal effect $\frac{\partial G(b)}{\partial b}$.

Application Decision. The application decision is forward-looking and individuals apply in period t , $\alpha_{i,t} = 1$, if

$$\Gamma(\theta_{i,t}, A_{i,t}, \chi_{i,t}) \equiv E \left[V_{i,t+1}^D - \left(\kappa_{i,t} V_{i,t+1}^E + (1 - \kappa_{i,t}) V_{i,t+1}^U \right) | \theta_{i,t}, A_{i,t}, \chi_{i,t} \right] \geq 0 \quad (\text{A.54})$$

where $\kappa_{i,t} = s_{i,t}$ if the individual is employed in period t and $\kappa_{i,t} = q_{i,t}$ if the individual is unemployed. At time t the value function of being employed is

$$V_{i,t}^E(X_{i,t}) = u(c_{i,t}^w) + \beta \cdot E \left[\alpha_{i,t} p_{i,t} V_{i,t+1}^D + (1 - \alpha_{i,t} p_{i,t}) \left[s_{i,t} V_{i,t+1}^E + (1 - s_{i,t}) V_{i,t+1}^U \right] | X_{i,t} \right] - \psi(\theta_{i,t}, e_{i,t}), \quad (\text{A.55})$$

where $s_{i,t} \equiv s(e_{i,t}; \Omega_t)$. The value function of being unemployed is

$$V_{i,t}^U(X_{i,t}) = u(c_{i,t}^z) + \beta \cdot E \left[\alpha_{i,t} p_{i,t} V_{i,t+1}^D + (1 - \alpha_{i,t} p_{i,t}) \left[q_{i,t} V_{i,t+1}^E + (1 - q_{i,t}) V_{i,t+1}^U \right] | X_{i,t} \right] - \phi(\theta_{i,t}, e_{i,t}), \quad (\text{A.56})$$

where $q_{i,t} \equiv q(e_{i,t}; \Omega_t)$. The value of being on DI, which is an absorbing state, is given by

$$V_{i,t}^D(X_{i,t}) = v(c_{i,t}^b) + \beta \cdot E \left[V_{i,t+1}^D | X_{i,t} \right]. \quad (\text{A.57})$$

For given state variables A_t and χ_t , the marginal applicant θ_t^A in period t is defined by

$$\Gamma(\theta_t^A, A_t, \chi_t) = 0. \quad (\text{A.58})$$

$\Gamma(\theta_t^A, A_t, \chi_t) = 0$ defines a unique marginal applicant for given A_t and χ_t if a disability shock today reduces the expected value of staying active in the labor market, i.e., $\frac{\partial E \left[\kappa_{i,t} V_{i,t+1}^E + (1 - \kappa_{i,t}) V_{i,t+1}^U | X_{i,t} \right]}{\partial \theta_t} < 0$. In this case θ_t^A conditional on A_t and χ_t is uniquely

defined by (A.58) because $E \left[V_{i,t+1}^D | X_{i,t} \right]$ does not depend on the disability level itself, i.e., $\frac{\partial E \left[V_{i,t+1}^D | X_{i,t} \right]}{\partial \theta_t} = 0$, while $\frac{\partial E \left[\kappa_{i,t} V_{i,t+1}^E + (1 - \kappa_{i,t}) V_{i,t+1}^U | X_{i,t} \right]}{\partial \theta_t} < 0$ and hence $E \left[V_{i,t+1}^D | X_{i,t} \right]$ and $E \left[\kappa_{i,t} V_{i,t+1}^E + (1 - \kappa_{i,t}) V_{i,t+1}^U | X_{i,t} \right]$ have single crossing.

A change in the DI benefit b has the following DI inflow effect at age $t + 1$:

$$\frac{\partial}{\partial b} DI_{t+1} = \frac{\partial}{\partial b} \int E \left[\int_{\theta_t^A}^{\infty} p(\theta) dF(\theta) | A_t, \chi_t \right] dG(A_t, \chi_t) \quad (\text{A.59})$$

$$= \int E \left[-\frac{\partial \theta_t^A}{\partial b} p(\theta_t^A) f(\theta_t^A) | A_t, \chi_t \right] dG(A_t, \chi_t), \quad (\text{A.60})$$

where $G(A_t, \chi_t)$ denotes the distribution of the state variables A_t and χ_t . Equation (A.60) sums up the responses from all marginal applicants. As in Appendix, A.3.1 we take the distribution $G(A_t, \chi_t)$ as given. Equation (A.58), the implicit function theorem, and the envelope theorem imply that

$$\frac{\partial \theta_t^A}{\partial b} = -\frac{\frac{\partial E \left[V_{i,t+1}^D - (\kappa_{i,t} V_{i,t+1}^E + (1 - \kappa_{i,t}) V_{i,t+1}^U) | \theta_{i,t}, A_{i,t}, \chi_{i,t} \right]}{\partial b}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.61})$$

$$= -\frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) | \theta_t^A, A_t, \chi_t \right] \quad (\text{A.62})$$

where $1 - DI_{f,s} \equiv \prod_{j=f}^s (1 - \alpha_j p_j)$. Combining (A.60) and (A.62) yields

$$\begin{aligned} \frac{\partial}{\partial b} DI_{t+1} &= \int E \left[c(\theta_t^A) \left(v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in Q^A \right], \end{aligned} \quad (\text{A.63})$$

where $c(\theta_t^A) \equiv \frac{p(\theta_t^A) f(\theta_t^A)}{\partial \Gamma_t / \partial \theta_t^A}$ and $\delta_{t+1}^k \equiv \beta^k (1 - DI_{t+1,t+k})$. Therefore, as in Appendix A.3.1 the inflow effect at age $t + 1$ corresponds to the (weighted) average marginal utility of consumption of all marginal applicants at age t when on DI benefits, which we denote by expression (A.63).

Analogously, a permanent change in wages implies

$$\frac{\partial \theta_t^A}{\partial w} = - \frac{\frac{\partial E \left[V_{i,t+1}^D - (\kappa_{i,t} V_{i,t+1}^E + (1-\kappa_{i,t}) V_{i,t+1}^U) \mid \theta_{i,t}, A_{i,t}, \chi_{i,t} \right]}{\partial w}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.64})$$

$$= \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k E_{t+k} u'(c_{t+1+k}^w) \mid \theta_t^A, A_t, \chi_t \right], \quad (\text{A.65})$$

where E_t denotes the share of employed in period $t + 1$. The inflow effect at age $t + 1$ is given by

$$\begin{aligned} \frac{\partial}{\partial w} DI_{t+1} &= - \int E \left[c(\theta_t^A) \left(E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k E_{t+k} u'(c_{t+1+k}^w) \right) \mid \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv - E \left[E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \omega_{t+1}^k u'(c_{t+1+k}^w) \mid \theta_t^A \in Q^A \right] \end{aligned} \quad (\text{A.66})$$

where $\omega_{t+1}^k \equiv \beta^k E_{t+k}$. Our empirical estimate of the DI inflow effect sums up the inflow effect at all ages, i.e.,

$$\frac{\partial DI / \partial b}{-\partial DI / \partial w} = \frac{\sum_{t=0}^{T-2} \frac{\partial}{\partial b} DI_{t+1}}{-\sum_{t=0}^{T-2} \frac{\partial}{\partial w} DI_{t+1}} \quad (\text{A.67})$$

$$= \frac{\sum_{t=0}^{T-2} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) \mid \theta_t^A \in Q^A \right]}{\sum_{t=0}^{T-2} E \left[E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \omega_{t+1}^k u'(c_{t+1+k}^w) \mid \theta_t^A \in Q^A \right]}. \quad (\text{A.68})$$

As in Appendix A.3.1, the ratio of inflow responses wrt. to benefits and wages is a ratio of the weighted average marginal utility of consumption when on DI and the weighted average marginal utility of consumption when working for marginal applicants. We again assume that the insurance value of marginal applicants provides a lower bound for the insurance value of DI benefits in the population.

Temporary Economic Shocks. We now consider temporary economic shocks Ω that affect wages and employment probabilities analogous to the static model from Appendix A.3.2. The temporary economic shocks $\Omega = \{d\Omega_k\}_{k=t}^{T-1}$ change the economic conditions in period k with intensity $d\Omega_k$ (this can also be zero, i.e., no effect in period k). The economic shock affects the wage, $\frac{\partial w_{i,t}}{\partial \Omega}$, the layoff probability, $\frac{\partial(1-s_{i,t})}{\partial \Omega}$, and the

job finding probability, $\frac{\partial q_{i,t}}{\partial \Omega}$. These temporary economic shocks Ω capture our empirical design where we use Bartik IVs as well as oil price shocks to leverage regional exposure to temporary economic shocks.

The effect of a temporary economic shock Ω on DI inflow is given by

$$\frac{\partial}{\partial \Omega} DI_{t+1} = \int E \left[-\frac{\partial \theta_t^A}{\partial \Omega} p(\theta_t^A) f(\theta_t^A) | A_t, \chi_t \right] dG(A_t, \chi_t), \quad (\text{A.69})$$

where

$$\frac{\partial \theta_t^A}{\partial \Omega} = -\frac{\frac{\partial E \left[V_{i,t+1}^D - (\kappa_{i,t} V_{i,t+1}^E + (1-\kappa_{i,t}) V_{i,t+1}^U) | \theta_{i,t}, A_{i,t}, \chi_{i,t} \right]}{\partial \Omega}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.70})$$

$$\begin{aligned} &= \left(\frac{\partial \Gamma_t}{\partial \theta_t^A} \right)^{-1} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[E_{i,t+k} u'(c_{i,t+1+k}^E) \frac{\partial w_{t+1+k}}{\partial \Omega} \right] | \theta_t^A, A_t, \chi_t \right] + \\ &\left(\frac{\partial \Gamma_t}{\partial \theta_t^A} \right)^{-1} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[\frac{\partial E_{i,t+k}}{\partial \Omega} \left[u(c_{i,t+1+k}^E) - \psi_{i,t+1+k} - \left(u(c_{i,t+1+k}^U) - \phi_{i,t+1+k} \right) \right] \right] | \theta_t^A, A_t, \chi_t \right], \end{aligned} \quad (\text{A.71})$$

and $\psi_{i,t} \equiv \psi(\theta_{i,t}, e_{i,t})$ and $\phi_{i,t} \equiv \phi(\theta_{i,t}, e_{i,t})$. $E_{i,t}$ denotes the probability to start out employed in period $t+1$ and $\frac{\partial E_{i,t+k}}{\partial \Omega}$ denotes the change in the employment probability due to the economic shock Ω .²⁰ If we bound the flow utility of job loss analogously to the static model in Appendix A.2, $u(c_{i,t+k}^E) - \psi_{i,t+k} - \left(u(c_{i,t+k}^U) - \phi_{i,t+k} \right) \geq u'(c_{i,t+k}^E) (w_{i,t+k} - z_{i,t+k})$, we have

$$\frac{1}{PVW_t} \frac{\partial \theta_t^A}{\partial \Omega} \geq \left(\frac{\partial \Gamma_t}{\partial \theta_t^A} \right)^{-1} \frac{1}{PVW_t} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[u'(c_{i,t+1+k}^E) \left(E_{i,t+k} \frac{\partial w_{t+1+k}}{\partial \Omega} + \frac{\partial E_{i,t+k}}{\partial \Omega} (w_{i,t+1+k} - z_{i,t+1+k}) \right) \right] | \theta_t^A, A_t, \chi_t \right],$$

where $PVW_t = E \left[\sum_{k=0}^{T-t-2} \beta^k \left[\left(E_{i,t+k} \frac{\partial w_{t+1+k}}{\partial \Omega} + \frac{\partial E_{i,t+k}}{\partial \Omega} (w_{i,t+1+k} - z_{i,t+1+k}) \right) \right] | \theta_t^A, A_t, \chi_t \right]$ is the present value effect of the economic shock on income. The ratio of the inflow response with respect to benefits and economics shocks rescaled by their present value

²⁰For illustration take an individual who is employed in period t . The probability to be employed in period $t+1$ is given by $E_{i,t} = s_{i,t}$, the probability to be employed in $t+2$ is $E_{i,t+1} = s_{i,t} \cdot s_{i,t+1} + (1-s_{i,t}) \cdot q_{i,t+1}$ and so on. An economic shock that starts in period t has the following effect on employment in $t+2$ for individuals being employed in t : $\frac{\partial E_{i,t+2}}{\partial \Omega} = \frac{\partial s_{i,t}}{\partial \Omega} \cdot (s_{i,t+1} - q_{i,t+1}) + s_{i,t} \cdot \frac{\partial s_{i,t+1}}{\partial \Omega} + (1-s_{i,t}) \cdot \frac{\partial q_{i,t+1}}{\partial \Omega}$.

impact identifies the ratio of marginal utilities of the marginal applicants

$$\frac{\partial DI / \partial b}{-\partial DI / \partial \Omega} = \frac{\frac{1}{PVB} \sum_{t=0}^{T-2} \frac{\partial}{\partial b} DI_{t+1}}{\frac{1}{PVW} \sum_{t=0}^{T-2} \frac{\partial}{\partial \Omega} DI_{t+1}} \quad (\text{A.72})$$

$$\begin{aligned} &\leq \frac{\sum_{t=0}^{T-2} \frac{1}{PVB_t} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) \mid \theta_t^A \in Q^A \right]}{\sum_{t=0}^{T-2} \frac{1}{PVW_t} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[u'(c_{i,t+1+k}^E) \left(E_{i,t+k} \frac{\partial w_{t+1+k}}{\partial \Omega} + \frac{\partial E_{i,t+k}}{\partial \Omega} (w_{i,t+1+k} - z_{i,t+1+k}) \right) \right] \mid \theta_t^A \in Q^A \right]} \\ &\equiv \frac{\mathbb{E} \left[v'(c^b) \mid \theta_t^A \in Q^A \right]}{\mathbb{E} \left[u'(c^w) \mid \theta_t^A \in Q^A \right]}. \end{aligned} \quad (\text{A.73})$$

This is again a lower bound for the insurance value if the marginal applicant values DI benefits less than the average DI recipient.

Table B.1: Summary Statistics, Benefit Generosity Sample

| | ROC | | Quebec | |
|---------------------------|-----------|-----------|-----------|-----------|
| | 1982-1986 | 1987-1991 | 1982-1986 | 1987-1991 |
| <i>A. Outcomes</i> | | | | |
| DI benefit receipt (in %) | 1.25 | 2.21 | 1.33 | 1.92 |
| Non-employment (in %) | 15.61 | 15.38 | 19.26 | 17.86 |
| DI benefits of recipients | 5,000 | 8,700 | 6,600 | 8,800 |
| | (3,300) | (5,700) | (4,800) | (4,900) |
| Earnings | 37,400 | 43,700 | 33,700 | 38,600 |
| | (41,600) | (60,100) | (42,100) | (43,300) |
| <i>B. Characteristics</i> | | | | |
| Share Female | 0.51 | 0.52 | 0.51 | 0.52 |
| Share Married | 0.66 | 0.70 | 0.62 | 0.62 |
| Age | 34.04 | 38.96 | 34.11 | 39.02 |
| | (8.98) | (8.90) | (8.75) | (8.67) |
| No. Kids | 0.28 | 0.25 | 0.28 | 0.22 |
| | (0.65) | (0.59) | (0.66) | (0.56) |
| No. Observations | 6,490,475 | 6,443,405 | 2,316,395 | 2,292,985 |

Notes: The table reports summary statistics for the Rest of Canada (ROC) and Quebec in the years before (1982-1986) and after the reform (1987-1992). Standard deviations of continuous variables in parentheses. Dollar amounts are rounded to \$100 following the vetting guidelines of the research data center.

B Additional Results for DI Benefit Generosity

Figure B.1: Raw Trends in Main Outcome Variables



Notes: The figures display the evolution of the main variables of interest over time in ROC (red dots) and in Quebec (blue triangles). Log disability benefits and log earnings are constructed as $\log(1 + x)$ where x are annual disability benefits or earnings.

C Additional Results for Wage Shocks

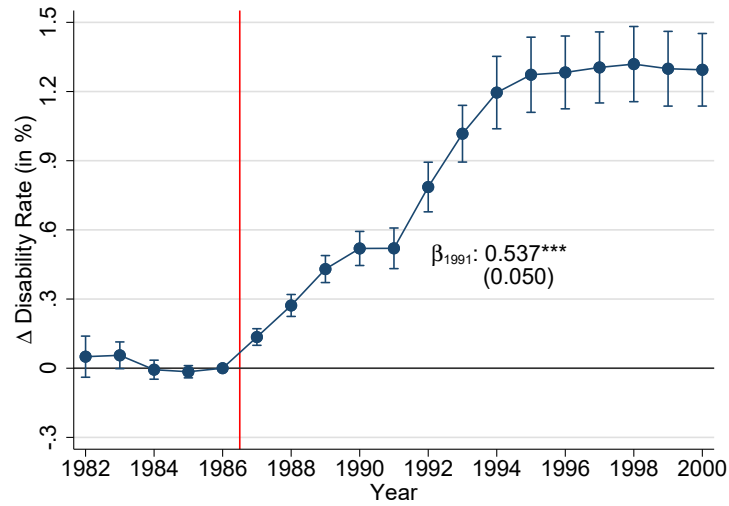
Table C.2: Impact of Local Labor Market Shocks on Employment

| | Pr(earnings>0) | Pr(earnings>10,000) | Pr(earnings>20,000) |
|-----------------|------------------|---------------------|---------------------|
| Coeff. estimate | -2.84 (-3.51) | 21.09*** (-4.86) | 54.79*** (-6.73) |
| Mean | 80.60 | 71.40 | 62.13 |
| Obs. | 18,829,205 | 18,829,205 | 18,829,205 |

Table C.3: The Impact of Earnings/Income on Disability Insurance Enrollment: Oil IV

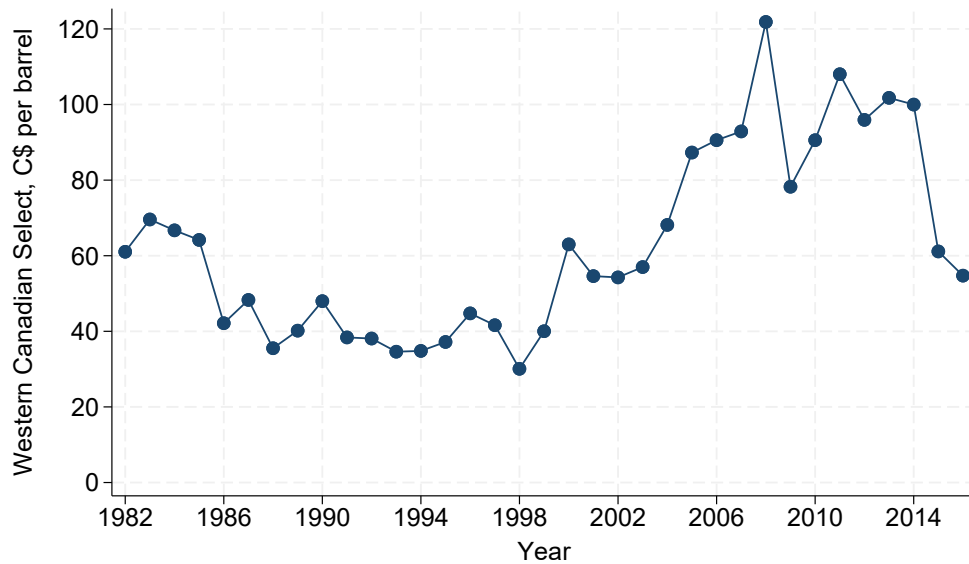
| | Earnings (\$1,000) | Current income (\$1,000) | Lifetime income (\$1,000) |
|--------------------------|----------------------|--------------------------|---------------------------|
| <i>A. Oil employment</i> | | | |
| Δ DI enrollment | -0.033*** (0.009) | -0.034*** (0.009) | -0.061*** (0.015) |
| 1st-stage coefficient | 212.0*** (17.6) | 210.0*** (17.4) | 117.0*** (9.3) |
| F-statistic | 145.9 | 145.3 | 156.3 |
| <i>B. Oil price</i> | | | |
| Δ DI enrollment | -0.022*** (0.002) | -0.023*** (0.002) | -0.072*** (0.007) |
| 1st-stage coefficients | | | |
| Current | 5.0*** (0.9) | 5.0*** (0.9) | 1.0*** (0.2) |
| 1st lag | 1.0 (1.8) | 1.0 (1.8) | 1.0*** (0.5) |
| 2nd lag | 3.0** (1.3) | 3.0** (1.3) | 0.0 (0.2) |
| F-statistic | 22.5 | 22.5 | 27.7 |
| Obs. | 18,829,205 | 18,829,205 | 18,829,205 |

Figure B.2: Difference in DI Take-Up Between RoC and Quebec by Year excluding Widows



Notes: The figure shows the estimated β_s -coefficients from the econometric specification in (12), excluding individuals whose spouse died in the same year to exclude survivor benefits. The capped spikes denote the upper and lower end of the 95-percent confidence interval.

Figure C.3: Oil Price over Time



Notes: The figure plots the real oil price trends between 1982 and 2016. We use Western Canadian Select (WCS) prices per barrel, published by the the Canadian Association of Petroleum Producers (<https://www.capp.ca/en/capp-data-centre/>).

Table C.4: OLS, Reduced-Form, and 2SLS Estimates for Different Outcomes

| | OLS | Reduced-Form | 2SLS |
|--------------------------|----------------------|----------------------|----------------------|
| Δ DI enrollment | -0.015*** (0.003) | -5.09*** (0.82) | -0.096*** (0.018) |
| Δ DI w/o survivor | -0.014*** (0.003) | -5.01*** (0.82) | -0.095*** (0.018) |
| log(DI benefit) | -0.001*** (0.000) | -0.457*** (0.078) | -0.009*** (0.002) |
| Pr(moving) | 0.003 (0.002) | -14.36 (31.12) | -0.271 (0.576) |
| Obs. | 18,829,205 | 18,829,205 | 18,829,205 |

Table C.5: Relation between Industry/Oil Shares and Census Division Characteristics

| | Industry shares | | | Oil shares |
|----------------------------|--------------------|-----------|----------------|------------|
| | Technical services | Farming | Transportation | |
| Share female | 0.095* | -0.376 | 0.119** | 0.001 |
| | (0.055) | (0.319) | (0.058) | (0.045) |
| Share married | -0.015 | 0.676*** | 0.021 | 0.047*** |
| | (0.018) | (0.133) | (0.030) | (0.017) |
| Family size | -0.014*** | -0.008 | -0.001 | -0.009** |
| | (0.003) | (0.024) | (0.004) | (0.004) |
| Age 15-19 | 0.062 | -0.350 | -0.092 | -0.160 |
| | (0.038) | (0.425) | (0.064) | (0.245) |
| Age 20-24 | -0.009 | -0.113 | -0.055 | -0.131 |
| | (0.046) | (0.373) | (0.056) | (0.254) |
| Age 25-29 | 0.002 | 0.181 | -0.217*** | -0.084 |
| | (0.051) | (0.351) | (0.069) | (0.248) |
| Age 30-34 | -0.107*** | -0.195 | -0.010 | -0.253 |
| | (0.035) | (0.340) | (0.057) | (0.231) |
| Age 35-39 | 0.054 | -1.559*** | 0.071 | -0.209 |
| | (0.040) | (0.446) | (0.077) | (0.259) |
| Age 40-44 | 0.067 | 0.381 | -0.171 | -0.228 |
| | (0.096) | (0.814) | (0.138) | (0.236) |
| Age 45-49 | 0.003 | -1.501* | -0.040 | -0.242 |
| | (0.076) | (0.831) | (0.097) | (0.257) |
| Age 50-54 | -0.021 | 0.283 | -0.063 | -0.312 |
| | (0.073) | (0.711) | (0.132) | (0.251) |
| Age 55-59 | 0.017 | 1.292** | -0.017 | -0.166 |
| | (0.104) | (0.597) | (0.112) | (0.250) |
| Pr(died by age 60) | 0.463 | -17.488 | -1.597 | 0.183 |
| | (1.618) | (12.056) | (1.940) | (1.577) |
| Share earnings 2. quartile | -0.046 | 0.620 | 0.039 | 0.012 |
| | (0.047) | (0.504) | (0.086) | (0.016) |
| Share earnings 3. quartile | -0.001 | 1.461** | 0.052 | 1.001 |
| | (0.069) | (0.583) | (0.130) | (0.856) |
| Share earnings 4. quartile | 0.109 | -0.680 | -0.007 | 0.581 |
| | (0.100) | (0.604) | (0.117) | (1.228) |
| Share income 2. quartile | 0.065 | -0.258 | -0.055 | 0.053 |
| | (0.048) | (0.585) | (0.097) | (0.328) |
| Share income 3. quartile | 0.030 | -1.514** | 0.007 | -0.458 |
| | (0.072) | (0.646) | (0.133) | (1.020) |
| Share income 4. quartile | -0.038 | 0.687 | 0.022 | -0.422 |
| | (0.106) | (0.665) | (0.133) | (1.173) |