

Supplemental Appendix

Measuring the Value of Disability Insurance from Take-Up Decisions

Andreas Haller, Norwegian School of Economics

Stefan Staubli, University of Calgary

A Theoretical Framework: Extensions and Robustness

This Appendix derives the optimal benefit formula from Section 2 in Appendix A.1 and discusses three extensions of the simple model from Section 2. The first extension in Appendix A.2 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 entrants. Appendix A.2 then also discusses how the insurance value of the marginal entrant relates to the insurance value of the average DI recipient. The second extension in Appendix A.3 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 (9) from exogenous labor market shocks that shift both wages and employment probabilities. The third extension in Appendix A.4 provides a dynamic model to address the question how we can make permanent DI benefit changes comparable to temporary wage changes.

A.1 Derivation Optimal DI Benefits

Combining equations (3)-(5) we can write the welfare effect of a balanced-budget increase of DI benefits as

$$\begin{aligned}\frac{dW(b)}{db} &= - \int_0^{\theta^A} \left[u'(w - \tau) \frac{d\tau}{db} \right] dF(\theta) + \int_{\theta^A}^{\infty} \left[p(\theta)v'(b) - (1 - p(\theta))(u'(w - \tau) \frac{d\tau}{db}) \right] dF(\theta) \\ &= - [1 - DI] u'(w - \tau) \frac{d\tau}{db} + DI \cdot v'(b)\end{aligned}\tag{A.1}$$

where $\frac{d\tau}{db} = \frac{DI}{1-DI} \left(1 + \frac{dDI}{db} \frac{b}{DI} \frac{1}{1-DI} \right) = \frac{DI}{1-DI} (1 + \frac{\varepsilon_{DI,b}}{1-DI})$ following from the balanced budget constraint and $\frac{dDI}{db} = \frac{d\theta^A}{db} p(\theta^A) f(\theta^A)$. Because of the envelope theorem changes in DI take-up ($\frac{d\theta^A}{db}$) have no first-order welfare effects and only enter through the government budget constraint ($\frac{d\tau}{db}$). Setting $\frac{dW(b)}{db} = 0$ for the optimal benefit level and rearranging yields the classic Baily-Chetty formula in equation (6).

A.2 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 entrant 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 entrants.

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 entrant 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

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

Thus, our approach still measures the insurance value of marginal entrants, 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 entrant for a given wage, DI benefit, and unearned income is given by

$$\theta^A(w, b, A) = u(c^w) - v(c^b). \quad (\text{A.3})$$

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,

$$-\frac{\partial DI}{\partial b} / \frac{\partial DI}{\partial w} = \frac{\int [v'(c^b) f(\theta^A) p(\theta^A) db] dG(w, b, A)}{\int [u'(c^w) f(\theta^A) p(\theta^A) dw] dG(w, b, A)}. \quad (\text{A.4})$$

The relative DI take-up responses with respect to DI benefits and wages still identify the insurance value of marginal entrants—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.4), an empirical challenge arises. The empirical variation 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 entrants for changes in benefits with those of marginal entrants for changes in wages. Reassuringly, Table 4 shows that the incidence of these two groups of marginal entrants 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 affect similar (w, A, θ) -types, suggesting we identify meaningful differences in marginal utilities and not composition effects.²⁶

²⁶Since DI benefits are a deterministic function of past earnings, we have comparable (w, A, b, θ) -types.

When is the Insurance Value of the Marginal Entrant a Lower Bound? A natural question is how representative is the marginal entrant's insurance value for the DI program's insurance value? What we want to measure is the value of a transfer between DI and non-DI recipients, which we can write as

$$IV_{pop} = \frac{E [v'(c^b(\theta); \theta) | \text{on DI}]}{E [u'(c^w(\theta); \theta) | \text{not on DI}]}, \quad (\text{A.5})$$

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 entrants

$$IV_A = \frac{E [v'(c^b(\theta); \theta) | \theta^A]}{E [u'(c^w(\theta); \theta) | \theta^A]}. \quad (\text{A.6})$$

The question is whether $IV_A \leq IV_{pop}$. We discuss this question in two steps. First, we analyze the case when the marginal utility of consumption only depends on consumption and is independent of the disability level θ . Second, we discuss the case when the marginal utility of consumption varies with the disability level.

If the marginal utility of consumption is independent of the disability level—the assumption in Diamond and Sheshinski (1995)—then $IV_A \leq IV_{pop}$ holds if (i) utility functions are concave, i.e., the marginal utility of consumption falls with consumption, (ii) the marginal entrant is weakly better off than the average DI recipient when on DI benefits, $E [c^b(\theta) | \theta^A] \geq E [c^b(\theta) | \text{on DI}]$, and (iii) the marginal entrant is weakly worse off than the average taxpayer when working, $E [c^w(\theta) | \theta^A] \leq E [c^w(\theta) | \text{not on DI}]$. Under assumptions (i)-(iii) we have

$$\frac{E [v'(c^b(\theta)) | \theta^A]}{E [u'(c^w(\theta)) | \theta^A]} \leq \frac{E [v'(c^b(\theta)) | \text{on DI}]}{E [u'(c^w(\theta)) | \text{not on DI}]} \quad (\text{A.7})$$

and the insurance value of marginal entrants indeed provides a lower bound for the true insurance value of DI benefits. Whether assumptions (ii) and (iii) hold is an empirical question. Consumption when on DI benefits is $c^b(\theta) = b(\theta) + A(\theta)$ – DI benefits plus other non-labor income. Consumption when working is $c^w(\theta) = w(\theta) + A(\theta)$ – labor income plus other non-labor income. We do not directly observe consumption in the

data. However, we can characterize the taxable incomes for the marginal entrant with our complier analysis from Section 5 and compare it to the average taxpayer and the average DI recipient.

Table 4 shows that marginal entrants are more likely to come from the lower end of the income distribution. Approximately 35% of all marginal entrants are in the lowest income quartile and only 10-15 percent are in the highest income quartile. Hence, marginal entrants are over-represented at lower incomes and under-represented at higher incomes (for equal representation, 25% of marginal entrants should be in each quartile). Income is measured three years before DI take-up, i.e., it is a measured when marginal entrants are still in the labor market and includes all types of taxable incomes, including labor and capital income. These patterns indicate that marginal entrants fare worse than the average taxpayer and support assumption (iii) $E [c^w(\theta)|\theta^A] \leq E [c^w(\theta)|\text{not on DI}]$.

Table B.3 compares marginal entrants and the average DI recipient.²⁷ Marginal entrants are on average older and less likely to be female. Interestingly, marginal entrants have a similar mortality as the average DI recipient. Marginal and average entrant have comparable incomes three years before DI take-up. Since average and marginal entrants have a very similar incidence across the income distribution, we might expect that they fare equally well when on DI benefits, i.e., $E [c^b(\theta)|\theta^A] \approx E [c^b(\theta)|\text{on DI}]$. Taken together, our complier analysis suggests that marginal entrants are similar to the average DI recipient in terms of income and mortality and less well off than the average taxpayer.

Even if assumptions (i)-(iii) hold, the marginal entrant's insurance value might not be a lower bound for the average insurance value if the marginal utility of consumption varies with the disability level. The marginal utility of consumption can vary with the disability level in two ways. The first scenario is that the marginal utility of consumption increases with the disability level, 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.

²⁷Tables C.7 and C.9 provide the same comparisons for the marginal entrant of the industry shift-share and the oil and gas instruments.

Hence, the marginal entrants' insurance value provides a lower bound in this scenario.

The second scenario is that the marginal utility of consumption is lower for higher disability levels. Then, the insurance value of the marginal entrants might not be a lower bound for the actual insurance value.

Therefore, an important determinant of the lower-bound result is how marginal utility varies with health. The empirical evidence on such state dependence is inconclusive. Some studies find that marginal utility declines in worse health states (Finkelstein, Luttmer and Notowidigdo, 2013), while others find the opposite (Lillard and Weiss, 1997). Additional evidence points to limited average differences but substantial heterogeneity across individuals (Brown, Goda and McGarry, 2016). As a result, whether the marginal entrant's insurance value is below that of the average DI recipient is ultimately an empirical question. The DI literature provides little direct evidence on this comparison. To our knowledge, no study compares the insurance value of marginal entrants with the average insurance value of the DI program, leaving this as an important area for future research.

A.3 Model with Three Labor Market States

Job Search Model. We now extend the model with a third labor market state—unemployed—and endogenous job search to address an empirical challenge. From a theoretical perspective, this extension does not change our results as long as the relevant outside option for the marginal entrant is work—an assumption we relax below. The relative DI responses to changes in benefits and wages continue to 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

$$\max_e s(e; \Omega) \cdot u(w(e; \Omega)) + (1 - s(e; \Omega)) \cdot v(z) - \psi(e; \theta) \quad (\text{A.8})$$

An interior solution fulfills the following first-order condition

$$s'(e; \Omega) \cdot [u(w(e; \Omega)) - v(z)] + s(e; \Omega) \cdot u'(w; \Omega) \cdot \frac{\partial w(e; \Omega)}{\partial e} - \psi'(e; \theta) = 0. \quad (\text{A.9})$$

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:

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

The marginal entrant θ^A is determined by

$$\Theta \equiv s(e; \Omega) \cdot u(w) + (1 - s(e; \Omega)) \cdot v(z) - \psi(e; \theta^A) - v(b) = 0. \quad (\text{A.11})$$

From (A.11), a change in DI benefits has the following impact on application behavior:

$$\frac{\partial \theta^A}{\partial b} = \frac{v'(b)}{\partial \Theta / \partial \theta^A}. \quad (\text{A.12})$$

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

$$\frac{\partial \theta^A}{\partial \Omega} = -\frac{1}{\partial \Theta / \partial \theta^A} \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [u(w) - v(z)] + s \cdot u'(w) \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]. \quad (\text{A.13})$$

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

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

$$\leq \frac{v'(b)}{u'(w) \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [w - z] + s \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]}. \quad (\text{A.15})$$

The inequality in (A.15) 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

$$\frac{\frac{\partial \theta^A}{\partial b}}{-\frac{\partial \theta^A}{\partial \Omega} / \left[\frac{\partial s(e; \Omega)}{\partial \Omega} [w - z] + s \cdot \frac{\partial w(e; \Omega)}{\partial \Omega} \right]} \leq \frac{v'(b; \theta^A)}{u'(w; \theta^A)} \leq \frac{E[v'(b; \theta) | \text{on DI}]}{E[u'(w; \theta) | \text{working}]} \quad (\text{A.16})$$

Hence, our empirical estimate that rescales the inflow effect by income is a lower bound for the insurance value of the marginal entrant.

Benefit Substitution Model. For some marginal entrants, the outside option is not necessarily work but may instead be other welfare benefits. In such cases, an increase in DI benefit generosity induces benefit substitution for a subset of marginal entrants. This extension shows that benefit substitution creates an upward bias in our estimator, as changes in DI benefits generate inflows from other welfare programs while changes in wages do not. The magnitude of this bias can be quantified empirically.

Such benefit substitution can occur in a model where individuals differ in their level of disability and application costs. We assume that there are other welfare benefits z and a DI application has cost ψ . An individual with disability severity θ and application costs ψ applies for DI benefits if their expected gain from receiving DI benefits exceeds

their application costs:

$$p(\theta) [v(b) - \max \{u(w - \tau) - \theta, v(z)\}] - \psi \geq 0. \quad (\text{A.17})$$

Individuals with high application costs might not apply for DI benefits even if they are severely disabled such that they do not work and claim other welfare benefits. In this setting there is a marginal entrant for each level of application costs, $\theta^A(\psi)$. The marginal entrants are implicitly defined by

$$p(\theta^A(\psi)) [v(b) - \max \{u(w - \tau) - \theta^A(\psi), v(z)\}] - \psi = 0. \quad (\text{A.18})$$

The relevant outside option to DI benefits differs by disability severity, $\max \{u(w - \tau) - \theta^A(\psi), v(z)\}$. Let $\theta^Z = u(w - \tau) - v(z)$ denote the cutoff for the outside option where individuals with $\theta > \theta^Z$ chose other benefits over working. In this model the share of individuals on DI is given by

$$DI = \int_0^\infty \int_{\theta^A(\psi)}^\infty p(\theta) dF(\theta|\psi) dH(\psi), \quad (\text{A.19})$$

where $F(\theta|\psi)$ is the distribution of disability severity conditional on a given level of application costs and $H(\cdot)$ denotes the distribution of application costs.

A change in DI benefit generosity then affects DI take-up by

$$\frac{\partial DI}{\partial b} = \int_0^\infty p(\theta^A) f(\theta^A|\psi) \left[1 \{ \theta^A > \theta^Z \} \frac{v'(b)}{\partial \tilde{F} / \partial \theta^A} + 1 \{ \theta^A \leq \theta^Z \} \frac{v'(b)}{\partial F / \partial \theta^A} \right] dH(\psi), \quad (\text{A.20})$$

which illustrates that take-up can come from either individuals on other benefits $\theta^A > \theta^Z$ or from individuals working $\theta^A \leq \theta^Z$.²⁸

A change in wages affects DI take-up by

$$\frac{\partial DI}{\partial w} = - \int_0^\infty p(\theta^A) f(\theta^A|\psi) \left[1 \{ \theta^A \leq \theta^Z \} \frac{u'(w - \tau)}{\partial F / \partial \theta^A} \right] dH(\psi) \quad (\text{A.21})$$

²⁸ $\tilde{F} \equiv p(\theta^A(\psi)) [v(b) - v(z)] - \psi$ and $F \equiv p(\theta^A(\psi)) [v(b) - \{u(w - \tau) - \theta^A(\psi)\}] - \psi$

and hence is only driven by individuals for whom the outside option is working.

In this setup, the ratio of take-up responses with respect to benefits and wages is

$$\begin{aligned} \frac{\partial DI / \partial b}{-\partial DI / \partial w} &= \frac{v'(b)}{u'(w - \tau)} + \frac{\int_0^\infty p(\theta^A) f(\theta^A | \psi) \left[1 \{ \theta^A \leq \theta^Z \} \frac{v'(b)}{\partial F / \partial \theta^A} \right] dH(\psi)}{\int_0^\infty p(\theta^A) f(\theta^A | \psi) \left[1 \{ \theta^A \leq \theta^Z \} \frac{u'(w - \tau)}{\partial F / \partial \theta^A} \right] dH(\psi)} \\ &= \frac{v'(b)}{u'(w - \tau)} + \frac{\partial DI / \partial b | \text{from other benefits}}{-\partial DI / \partial w}. \end{aligned} \quad (\text{A.22})$$

A key implication is that wage changes only affect individuals whose outside option is work, whereas DI benefit changes also affect individuals whose outside option is other welfare programs. In this setting, our revealed preference method identifies the insurance value, $\frac{v'(b)}{u'(w - \tau)}$, plus a bias term that reflects the extent to which increases in DI generosity induce inflows from other welfare programs relative to those induced by wage changes. As a result, our revealed preference method overstates the insurance value, but the magnitude of this bias can be quantified empirically by decomposing inflows into transitions from work and from other benefits.

A.4 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.4.1. We then discuss the dynamic version of the model from Appendix A.3 with three labor market states in Section A.4.2. 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.4.1 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

$$A_{i,t+1} = (1 + r_t)A_{i,t} + b(X_{i,t}) - c_{i,t}^b(X_{i,t}) \quad (\text{A.23})$$

$$A_{i,t+1} = (1 + r_t)A_{i,t} + w(X_{i,t}) - \tau_{i,t}(X_{i,t}) - c_{i,t}^w(X_{i,t}), \quad (\text{A.24})$$

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 $t + 1$:

$$DI_{i,t+1} = 1 - \left[\prod_{k=0}^t (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k})) \right]. \quad (\text{A.25})$$

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, disability level $\theta_{i,t}$ 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.²⁹ The agent's problem is given by

$$\begin{aligned}
V_i(b) = \max E & \left[\sum_{t=0}^{T-1} \beta^t \left(v(c_{i,t}^b) \cdot D_{i,t} + \left(u(c_{i,t}^w) - \psi(\theta_{i,t}) \right) \cdot (1 - D_{i,t}) \right) \right] \\
& + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \left((1 + r_t) A_{i,t} + b_{i,t} - c_{i,t}^b - A_{i,t+1} \right) D_{i,t} \right] \\
& + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^W \left((1 + r_t) A_{i,t} + w_{i,t} - \tau_{i,t} - c_{i,t}^w - A_{i,t+1} \right) (1 - D_{i,t}) \right]
\end{aligned} \tag{A.26}$$

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

$$\max_b W(b) = \int_i V_i(b) di + \lambda \cdot G(b) \tag{A.27}$$

where

$$G(b) = \int_i E \left[\sum_{t=0}^{T-1} (1 + r_t)^{-t} \left((1 - D_{i,t}) \cdot \tau_{i,t} - D_{i,t} \cdot b_{i,t} \right) \right] di \tag{A.28}$$

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

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.³⁰ The optimal DI benefit solves

$$\frac{\partial W(b)}{\partial b} = E \left[\int_i \sum_{t=0}^{T-1} \beta^t \left(v'(c_{i,t}^b) \cdot D_{i,t} \right) di \right] + \lambda \cdot \frac{\partial G(b)}{\partial b} = 0 \tag{A.29}$$

²⁹The 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.

³⁰The 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 significant 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.

where the budget effect is

$$\frac{\partial G(b)}{\partial b} = - \int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} \left(D_{i,t} + \frac{\partial D_{i,t}}{\partial b} \cdot (b + \tau) \right) \right] di \quad (\text{A.30})$$

$$= -DI \left[1 + \frac{\partial DI}{\partial b} \frac{1}{DI} (b + \tau) \right] \quad (\text{A.31})$$

$$= -DI \left[1 + \frac{\partial DI}{\partial b} \frac{b}{DI} \frac{1}{1-DI} \right] \quad (\text{A.32})$$

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.³¹ We assume that the planner's budget constraint is differentiable, i.e., $\frac{\partial G(b)}{\partial b}$ exists.

Proof. of (A.29).

The first order condition in (A.29) 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 $U'(b) = 0$. 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

³¹Equation (A.32) uses the balanced budget constraint, implying $\tau = \frac{DI}{1-DI} \cdot b$.

function for $W(b)$, which implies that

$$\frac{\partial W(b)}{\partial b} = \frac{\partial L(b)}{\partial b} \quad (\text{A.33})$$

$$= E \left[\int_i \sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \cdot D_{i,t} di \right] + \lambda \cdot \frac{\partial G(b)}{\partial b} \quad (\text{A.34})$$

$$= E \left[\int_i \sum_{t=0}^{T-1} \beta^t \left(v'(c_{i,t}^b) \cdot D_{i,t} \right) di \right] + \lambda \cdot \frac{\partial G(b)}{\partial b} \quad (\text{A.35})$$

$$= \frac{\partial U(b)}{\partial b} = 0 \quad (\text{A.36})$$

and concludes the proof of (A.29). $\square$

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.29) at the Baily-Chetty formula

$$\frac{\mathbb{E} [v'(c^b) | DI = 1]}{\mathbb{E} [u'(c^w) | DI = 0]} = 1 + \frac{\varepsilon_{DI,b}}{1-DI}, \quad (\text{A.37})$$

where $\mathbb{E} [v'(c^b) | DI = 1] \equiv \frac{\int_i E \left[\sum_{t=0}^{T-1} \beta^t \cdot v'(c_{i,t}^b) \cdot D_{i,t} \right] di}{\int_i E \left[\sum_{t=0}^{T-1} (1+r_t)^{-t} (D_{i,t}) \right] di}$ 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

$$V_{i,t}^E (X_{i,t}) = u(c_{i,t}^w) - \psi(\theta_{i,t}) + \beta \cdot E \left[\alpha_{i,t} p_{i,t} V_{i,t+1}^D + (1 - \alpha_{i,t} p_{i,t}) V_{i,t+1}^E | X_{i,t} \right]. \quad (\text{A.38})$$

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

$$V_{i,t}^D (X_{i,t}) = v(c_{i,t}^b) + \beta \cdot E \left[V_{i,t+1}^D | X_{i,t} \right]. \quad (\text{A.39})$$

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:

$$\Gamma(\theta_{i,t}, A_{i,t}, \chi_{i,t}) \equiv E[V_{i,t+1}^D - V_{i,t+1}^E | \theta_{i,t}, A_{i,t}, \chi_{i,t}] \geq 0. \quad (\text{A.40})$$

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

$$\Gamma(\theta_t^A, A_t, \chi_t) = 0. \quad (\text{A.41})$$

$\Gamma(\theta_t^A, A_t, \chi_t) = 0$ defines a unique marginal entrant 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., $\frac{\partial E[V_{i,t+1}^E | X_{i,t}]}{\partial \theta_t} < 0$.³² In this case, θ_t^A conditional on A_t and χ_t is uniquely defined by (A.41), 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.41) implicitly defines the marginal entrant 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 the expected value of working. θ_t^A is then differentiable in b by the implicit function

³²Persistence 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.

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

$$\frac{\partial}{\partial b} DI_{t+1} = \frac{\partial}{\partial b} \int E \left[\int_{\theta_t^A}^{\infty} p(\theta) dF(\theta) | A_t, \chi_t \right] dG(A_t, \chi_t) \quad (\text{A.42})$$

$$= \int E \left[-\frac{\partial \theta_t^A}{\partial b} p(\theta_t^A) f(\theta_t^A) | A_t, \chi_t \right] dG(A_t, \chi_t), \quad (\text{A.43})$$

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 entrants so that different disability cutoffs θ_t^A apply for different state variables A_t and χ_t . Equation (A.43) sums up the responses from all marginal entrants. Furthermore, we take the distribution $G(A_t, \chi_t)$ as given, which means that equation (A.43) 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. Equation (A.41), the implicit function theorem, and the envelope theorem imply that

$$\frac{\partial \theta_t^A}{\partial b} = - \frac{\frac{\partial E[V_{i,t+1}^D - V_{i,t+1}^E | \theta_t^A, A_t, \chi_t]}{\partial b}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.44})$$

$$= -\frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) | \theta_t^A, A_t, \chi_t \right] \quad (\text{A.45})$$

where $1 - DI_{f,s} \equiv \prod_{j=f}^s (1 - \alpha_j p_j)$. Combining (A.43) and (A.45) yields

$$\begin{aligned} \frac{\partial}{\partial b} DI_{t+1} &= \int E \left[h(\theta_t^A) \left(v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in \mathcal{Q}^A \right], \end{aligned} \quad (\text{A.46})$$

where $h(\theta_t^A) \equiv \frac{p(\theta_t^A) f(\theta_t^A)}{\partial \Gamma_t / \partial \theta_t^A}$, 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 $(h(\theta_t^A))$ -weighted average marginal utility of consumption of all marginal entrants at age t when on DI benefits, which we denote by expression (A.46).

Analogously, a permanent change in wages implies

$$\frac{\partial \theta_t^A}{\partial w} = - \frac{\frac{\partial E[V_{i,t+1}^D - V_{i,t+1}^E | \theta_t^A, A_t, \chi_t]}{\partial w}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.47})$$

$$= \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) u'(c_{t+1+k}^w) | \theta_t^A, A_t, \chi_t \right] \quad (\text{A.48})$$

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

$$\begin{aligned} \frac{\partial}{\partial w} DI_{t+1} &= - \int E \left[h(\theta_t^A) \left(u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) u'(c_{t+1+k}^w) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv -E \left[u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k u'(c_{t+1+k}^w) | \theta_t^A \in Q^A \right]. \end{aligned} \quad (\text{A.49})$$

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

$$\frac{\partial DI / \partial b}{-\partial DI / \partial w} = \frac{\sum_{t=0}^{T-2} \frac{\partial}{\partial b} DI_{t+1}}{-\sum_{t=0}^{T-2} \frac{\partial}{\partial w} DI_{t+1}} \quad (\text{A.50})$$

$$= \frac{\sum_{t=0}^{T-2} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in Q^A \right]}{\sum_{t=0}^{T-2} E \left[u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k u'(c_{t+1+k}^w) | \theta_t^A \in Q^A \right]} \quad (\text{A.51})$$

Equation (A.51) indicates that the ratio of DI inflow responses with respect to changes in benefits and wages corresponds to the marginal entrants' 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 entrants' insurance value is lower in all periods than that of all DI recipients (as we assume in the static model).³³

³³We can use the mediant theorem to show this statement formally. Equation (A.51) is a sum of the form $\frac{\sum v_j}{\sum u_j}$, where $v_j \equiv h(\theta_{t(j)}^A) \delta_{t(j)+1}^k v'(c_{t(j)+1+k(j)}^b)$ and $u_j \equiv h(\theta_{t(j)}^A) \delta_{t(j)+1}^k u'(c_{t(j)+1+k(j)}^w)$ for all t, k -combinations in (A.51) for all marginal entrants (including $k = 0$, where $v_j \equiv h(\theta_{t(j)}^A) v'(c_{t(j)+1}^b)$ and $u_j \equiv h(\theta_{t(j)}^A) u'(c_{t(j)+1}^w)$). Let $m_j = \frac{v_j}{u_j} = \frac{v'(c_{t(j)}^b)}{u'(c_{t(j)}^w)}$ and the indexing j is such that $m_1 \leq \dots \leq m_j \leq \dots \leq m_N$. By the mediant theorem, $\frac{\partial DI / \partial b}{\partial DI / \partial w} \leq m_N$, and so if the marginal entrants' insurance value is always lower than that of all DI recipients, we have $\frac{\partial DI / \partial b}{\partial DI / \partial w} \leq m_N \leq \frac{\mathbb{E}[v'(c^b) | DI=1]}{\mathbb{E}[u'(c^w) | DI=0]}$.

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

$$\begin{aligned} \frac{\partial \theta_t^A}{\partial w_{T,t+1}} &= - \frac{\frac{\partial E[V_{i,t+1}^D - V_{i,t+1}^E | \theta_t^A, A_t, \chi_t]}{\partial w_{T,t+1}}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \\ &\approx \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[u'(c_{t+1}^w) dw_{i,t+1} + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) u'(c_{t+1+k}^w) dw_{i,t+1+k} | \theta_t^A, A_t, \chi_t \right]. \end{aligned} \quad (\text{A.52})$$

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, $\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T}$, compares to the effect of a permanent wage shock, $\frac{\partial \theta_t^A}{\partial w}$. Both $\frac{1}{PVW_{t+1}} \frac{\partial \theta_t^A}{\partial w_T}$ and $\frac{\partial \theta_t^A}{\partial w}$ measure weighted average marginal utilities of consumption of the marginal entrants. The difference 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 entrants. 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 entrants 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.4.2 Dynamic Model with Three Labor Market States.

Setup. We now extend the model with three labor market states from Appendix A.3 to a dynamic setting. The same notation, timing, and constraints from the dynamic model in Appendix A.4.1 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

$$A_{i,t+1} = (1 + r_t)A_{i,t} + b(X_{i,t}) - c_{i,t}^b(X_{i,t}) \quad (\text{A.53})$$

$$A_{i,t+1} = (1 + r_t)A_{i,t} + w(X_{i,t}) - \tau_{i,t}(X_{i,t}) - c_{i,t}^w(X_{i,t}) \quad (\text{A.54})$$

$$A_{i,t+1} = (1 + r_t)A_{i,t} + z(X_{i,t}) - c_{i,t}^z(X_{i,t}) \quad (\text{A.55})$$

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, individuals 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

$$DI_{i,t+1} = 1 - \left[\prod_{k=0}^t (1 - \alpha_{i,k}(X_{i,k}) \cdot p(\theta_{i,k})) \right]. \quad (\text{A.56})$$

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

$$E_{i,t+1} = (1 - DI_{i,t+1}) (s(e_{i,t}; \Omega_t) + q(e_{i,t}; \Omega_t)). \quad (\text{A.57})$$

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

$$\begin{aligned}
V_i(b) = \max E & \left[\sum_{t=0}^{T-1} \beta^t \left(v(c_{i,t}^b) \cdot D_{i,t} + \left(u(c_{i,t}^w) - \psi(\theta_{i,t}, e_{i,t}) \right) \cdot E_{i,t} + \left(u(c_{i,t}^z) - \phi(\theta_{i,t}, e_{i,t}) \right) \cdot U_{i,t} \right) \right] \\
& + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^D \left((1 + r_t) A_{i,t} + b_{i,t} - c_{i,t}^b - A_{i,t+1} \right) D_{i,t} \right] \\
& + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^W \left((1 + r_t) A_{i,t} + w_{i,t} - \tau_{i,t} - c_{i,t}^w - A_{i,t+1} \right) E_{i,t} \right] \\
& + E \left[\sum_{t=0}^{T-1} \beta^t \mu_{i,t}^Z \left((1 + r_t) A_{i,t} + z_{i,t} - c_{i,t}^z - A_{i,t+1} \right) U_{i,t} \right].
\end{aligned} \tag{A.58}$$

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

$$\max_b W(b) = \int_i V_i(b) di + \lambda \cdot G(b) \tag{A.59}$$

where

$$G(b) = \int_i E \left[\sum_{t=0}^{T-1} (1 + r_t)^{-t} (E_{i,t} \cdot \tau_{i,t} - D_{i,t} \cdot b_{i,t} - U_{i,t} \cdot z_{i,t}) \right] di \tag{A.60}$$

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.29). 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

$$\Gamma(\theta_{i,t}, A_{i,t}, \chi_{i,t}) \equiv E \left[V_{i,t+1}^D - \left(\kappa_{i,t} V_{i,t+1}^E + (1 - \kappa_{i,t}) V_{i,t+1}^U \right) | \theta_{i,t}, A_{i,t}, \chi_{i,t} \right] \geq 0 \tag{A.61}$$

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

$$\begin{aligned}
V_{i,t}^E(X_{i,t}) &= u(c_{i,t}^w) - \psi(\theta_{i,t}, e_{i,t}) \dots \\
&+ \beta \cdot E \left[\alpha_{i,t} p_{i,t} V_{i,t+1}^D + (1 - \alpha_{i,t} p_{i,t}) \left[s_{i,t} V_{i,t+1}^E + (1 - s_{i,t}) V_{i,t+1}^U \right] | X_{i,t} \right],
\end{aligned} \tag{A.62}$$

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

$$\begin{aligned} V_{i,t}^U(X_{i,t}) &= u(c_{i,t}^z) - \phi(\theta_{i,t}, e_{i,t}) \dots \\ &+ \beta \cdot E \left[\alpha_{i,t} p_{i,t} V_{i,t+1}^D + (1 - \alpha_{i,t} p_{i,t}) \left[q_{i,t} V_{i,t+1}^E + (1 - q_{i,t}) V_{i,t+1}^U \right] | X_{i,t} \right], \end{aligned} \quad (\text{A.63})$$

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

$$V_{i,t}^D(X_{i,t}) = v(c_{i,t}^b) + \beta \cdot E \left[V_{i,t+1}^D | X_{i,t} \right]. \quad (\text{A.64})$$

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

$$\Gamma(\theta_t^A, A_t, \chi_t) = 0. \quad (\text{A.65})$$

$\Gamma(\theta_t^A, A_t, \chi_t) = 0$ defines a unique marginal entrant 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.65) 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 $t + 1$:

$$\frac{\partial}{\partial b} DI_{t+1} = \frac{\partial}{\partial b} \int E \left[\int_{\theta_t^A}^{\infty} p(\theta) dF(\theta) | A_t, \chi_t \right] dG(A_t, \chi_t) \quad (\text{A.66})$$

$$= \int E \left[-\frac{\partial \theta_t^A}{\partial b} p(\theta_t^A) f(\theta_t^A) | A_t, \chi_t \right] dG(A_t, \chi_t), \quad (\text{A.67})$$

where $G(A_t, \chi_t)$ denotes the distribution of the state variables A_t and χ_t . Equation (A.67) sums up the responses from all marginal entrants. As in Appendix, A.4.1 we take the distribution $G(A_t, \chi_t)$ as given. Equation (A.65), the implicit function theorem,

and the envelope theorem imply that

$$\frac{\partial \theta_t^A}{\partial b} = - \frac{\frac{\partial E[V_{i,t+1}^D - (\kappa_{i,t} V_{i,t+1}^E + (1-\kappa_{i,t}) V_{i,t+1}^U) | \theta_{i,t}, A_{i,t}, \chi_{i,t}]}{\partial b}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.68})$$

$$= - \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) | \theta_t^A, A_t, \chi_t \right], \quad (\text{A.69})$$

where $1 - DI_{f,s} \equiv \prod_{j=f}^s (1 - \alpha_j p_j)$. Combining (A.67) and (A.69) yields

$$\begin{aligned} \frac{\partial}{\partial b} DI_{t+1} &= \int E \left[h(\theta_t^A) \left(v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \beta^k (1 - DI_{t+1,t+k}) v'(c_{t+1+k}^b) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in Q^A \right], \end{aligned} \quad (\text{A.70})$$

where $h(\theta_t^A) \equiv \frac{p(\theta_t^A) f(\theta_t^A)}{\partial \Gamma_t / \partial \theta_t^A}$ and $\delta_{t+1}^k \equiv \beta^k (1 - DI_{t+1,t+k})$. Therefore, as in Appendix A.4.1 the inflow effect at age $t + 1$ corresponds to the (weighted) average marginal utility of consumption of all marginal entrants at age t when on DI benefits, which we denote by expression (A.70).

Analogously, a permanent change in wages implies

$$\frac{\partial \theta_t^A}{\partial w} = - \frac{\frac{\partial E[V_{i,t+1}^D - (\kappa_{i,t} V_{i,t+1}^E + (1-\kappa_{i,t}) V_{i,t+1}^U) | \theta_{i,t}, A_{i,t}, \chi_{i,t}]}{\partial w}}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} \quad (\text{A.71})$$

$$= \frac{1}{\frac{\partial \Gamma_t}{\partial \theta_t^A}} E \left[E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k E_{t+k} u'(c_{t+1+k}^w) | \theta_t^A, A_t, \chi_t \right], \quad (\text{A.72})$$

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

$$\begin{aligned} \frac{\partial}{\partial w} DI_{t+1} &= - \int E \left[h(\theta_t^A) \left(E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \beta^k E_{t+k} u'(c_{t+1+k}^w) \right) | \theta_t^A, A_t, \chi_t \right] dM(\theta_t^A, A_t, \chi_t) \\ &\equiv -E \left[E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \omega_{t+1}^k u'(c_{t+1+k}^w) | \theta_t^A \in Q^A \right] \end{aligned} \quad (\text{A.73})$$

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.,

$$\frac{\partial DI/\partial b}{-\partial DI/\partial w} = \frac{\sum_{t=0}^{T-2} \frac{\partial}{\partial b} DI_{t+1}}{-\sum_{t=0}^{T-2} \frac{\partial}{\partial w} DI_{t+1}} \quad (\text{A.74})$$

$$= \frac{\sum_{t=0}^{T-2} E \left[v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in Q^A \right]}{\sum_{t=0}^{T-2} E \left[E_t u'(c_{t+1}^w) + \sum_{k=1}^{T-t-2} \omega_{t+1}^k u'(c_{t+1+k}^w) | \theta_t^A \in Q^A \right]}. \quad (\text{A.75})$$

As in Appendix A.4.1, the ratio of inflow responses with respect 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 entrants. We again assume that the insurance value of marginal entrants 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.4.2. 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

$$\frac{\partial}{\partial \Omega} DI_{t+1} = \int E \left[-\frac{\partial \theta_t^A}{\partial \Omega} p(\theta_t^A) f(\theta_t^A) | A_t, \chi_t \right] dG(A_t, \chi_t), \quad (\text{A.76})$$

where

$$\frac{\partial \theta_t^A}{\partial \Omega} = - \frac{\frac{\partial E[V_{i,t+1}^D - (\kappa_{i,t} V_{i,t+1}^E + (1-\kappa_{i,t}) V_{i,t+1}^U) | \theta_{i,t}^A, A_{i,t}, \chi_{i,t}]}{\partial \Omega}}{\frac{\partial \Gamma_t^A}{\partial \theta_t^A}} \quad (\text{A.77})$$

$$= \left(\frac{\partial \Gamma_t^A}{\partial \theta_t^A} \right)^{-1} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[E_{i,t+k} u'(c_{i,t+1+k}^E) \frac{\partial w_{t+1+k}}{\partial \Omega} \right] | \theta_t^A, A_t, \chi_t \right] + \left(\frac{\partial \Gamma_t^A}{\partial \theta_t^A} \right)^{-1} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[\frac{\partial E_{i,t+k}}{\partial \Omega} \left[u(c_{i,t+1+k}^E) - \psi_{i,t+1+k} - (u(c_{i,t+1+k}^U) - \phi_{i,t+1+k}) \right] \right] | \theta_t^A, A_t, \chi_t \right], \quad (\text{A.78})$$

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 .³⁴ If we bound the flow utility of job loss analogously to the static model in Appendix A.3, $u(c_{i,t+k}^E) - \psi_{i,t+k} - (u(c_{i,t+k}^U) - \phi_{i,t+k}) \geq u'(c_{i,t+k}^E) (w_{i,t+k} - z_{i,t+k})$, 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 impact identifies the ratio of marginal utilities of the marginal entrants

$$\frac{\partial DI / \partial b}{-\partial DI / \partial \Omega} = \frac{\frac{1}{PVB} \sum_{t=0}^{T-2} \frac{\partial}{\partial b} DI_{t+1}}{\frac{1}{PVW} \sum_{t=0}^{T-2} \frac{\partial}{\partial \Omega} DI_{t+1}} \quad (\text{A.79})$$

$$\begin{aligned} &\leq \frac{\sum_{t=0}^{T-2} \frac{1}{PVB_t} E [v'(c_{t+1}^b) + \sum_{k=1}^{T-t-2} \delta_{t+1}^k v'(c_{t+1+k}^b) | \theta_t^A \in \mathcal{Q}^A]}{\sum_{t=0}^{T-2} \frac{1}{PVW_t} E \left[\sum_{k=0}^{T-t-2} \beta^k \left[u'(c_{i,t+1+k}^E) \left(E_{i,t+k} \frac{\partial w_{t+1+k}}{\partial \Omega} + \frac{\partial E_{i,t+k}}{\partial \Omega} (w_{i,t+1+k} - z_{i,t+1+k}) \right) \right] | \theta_t^A \in \mathcal{Q}^A \right]} \\ &\equiv \frac{\mathbb{E} [v'(c^b) | \theta_t^A \in \mathcal{Q}^A]}{\mathbb{E} [u'(c^w) | \theta_t^A \in \mathcal{Q}^A]}. \end{aligned} \quad (\text{A.80})$$

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

³⁴For 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 : $\frac{\partial E_{i,t+2}}{\partial \Omega} = \frac{\partial s_{i,t}}{\partial \Omega} \cdot (s_{i,t+1} - q_{i,t+1}) + s_{i,t} \cdot \frac{\partial s_{i,t+1}}{\partial \Omega} + (1 - s_{i,t}) \cdot \frac{\partial q_{i,t+1}}{\partial \Omega}$.

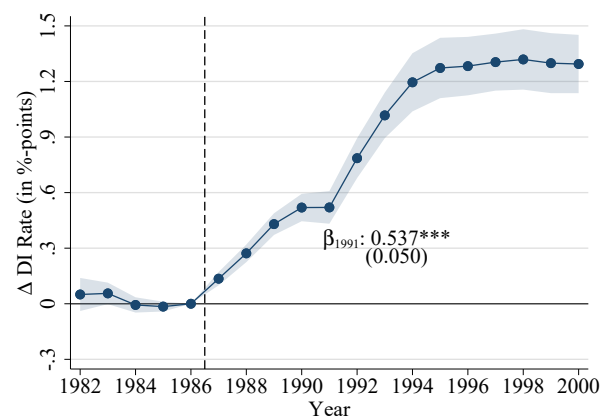
B Additional Results for DI Benefit Generosity

Table B.1: Summary Statistics, Benefit Generosity Sample

	ROC		Quebec	
	1982-1986	1987-1992	1982-1986	1987-1992
<i>A. Outcomes</i>				
DI benefit receipt (in %)	1.25	2.40	1.33	2.02
Non-employment (in %)	15.61	15.84	19.26	18.24
DI benefits of recipients	5,000 (3,400)	9,000 (5,600)	6,700 (4,800)	9,000 (4,500)
Earnings	16,700 (18,600)	24,600 (34,000)	15,000 (15,100)	21,700 (24,500)
<i>B. Characteristics</i>				
Share Female	0.51	0.52	0.51	0.52
Share Married	0.66	0.71	0.62	0.63
Age	34.05 (8.98)	39.45 (8.95)	34.11 (8.75)	39.51 (8.72)
No. Kids	0.28 (0.65)	0.25 (0.59)	0.28 (0.66)	0.22 (0.56)
No. Observations	6,491,185	7,728,240	2,315,605	2,747,770

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). All monetary outcomes are expressed in 2019 dollars using the national CPI. Standard deviations of continuous variables in parentheses. Dollar amounts are rounded to \$100 following the vetting guidelines of the research data center.

Figure B.1: Difference in DI Take-Up Between RoC and Quebec by Year excluding Widows



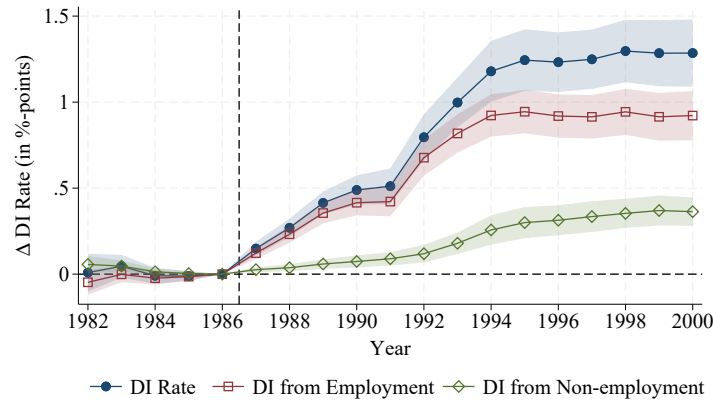
Notes: The figure shows the estimated β_s -coefficients from the econometric specification in (10), excluding individuals whose spouse died in the same year to exclude survivor benefits. The shaded area represents the 95-percent confidence interval.

Table B.2: Impacts of 1987 Reform on DI Receipt, Robustness

	Main estimate	Oil and gas provinces	Unbalanced sample
Coeff. estimate (β_{1991})	0.527*** (0.061)	0.404*** (0.056)	0.534*** (0.047)
Mean ($\bar{y}$)	1.84	1.59	2.15
Elasticity ($\beta_{1991} \cdot \frac{\bar{y}}{\bar{y}}$)	0.580*** (0.067)	0.507*** (0.070)	0.398*** (0.035)
No. Observations	17,543,255	9,232,205	28,433,095

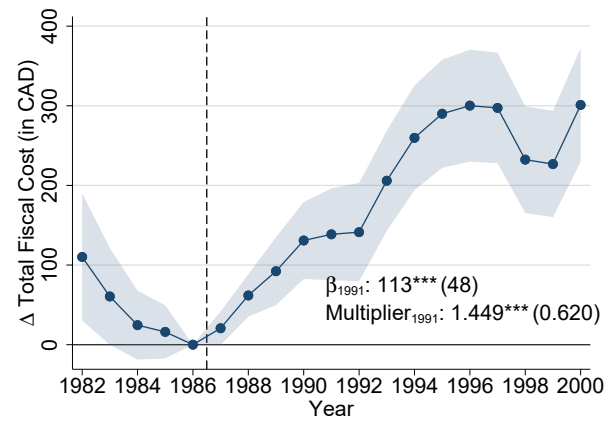
Notes: The table reports the reform effect in the fifth year after the reform for DI receipt (the β_{1991} -coefficient from equation 10). Coefficient estimates and standard errors are multiplied by 100 and should be interpreted as percentage points. The main sample includes 15-59 year-old individuals who filed taxes continuously from 1982 until 1991 or until death, whichever occurred first. Oil and gas provinces sample restricts the main sample to individuals residing in Alberta, British Columbia, Saskatchewan, Newfoundland and Labrador, and Quebec. The unbalanced sample includes all 15-59 year-old individuals from 1982 until 1991, regardless of whether they filed taxes continuously. Standard errors are reported in parentheses and are clustered at the Census Division. 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%.

Figure B.2: DI receipt from employment and non-employment



Notes: The figure reports the estimated β_s -coefficients from equation (10) for different outcomes. “DI Rate” is an indicator equal to one if an individual receives a public disability pension. “DI from Employment” is an indicator equal to one if the individual receives a public disability pension and had positive labor income in the two years before entering DI. “DI from Non-Employment” is an indicator equal to one if the individual receives a public disability pension and had no labor income in the two years before entering DI. The shaded area represents the 95-percent confidence interval.

Figure B.3: Total fiscal cost and fiscal multiplier



Notes: The figure reports the estimated β_s -coefficients from equation (10) for total fiscal cost, defined as DI benefit expenditures minus labor income tax revenues. The shaded area represents the 95-percent confidence interval.

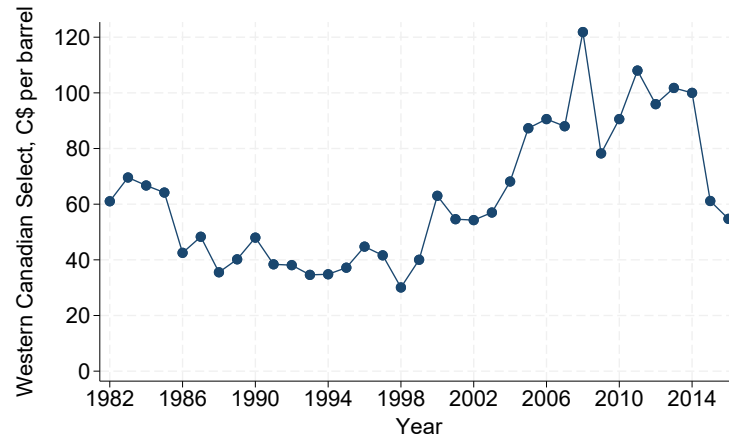
Table B.3: Comparison Marginal vs. Average Entrant, Benefit Change

	Marginal entrant	Average entrant	P-value
Age	50.57*** (0.81)	45.66*** (0.13)	0.000
Family size	3.60*** (0.26)	3.04*** (0.02)	0.027
Share female	0.426*** (0.038)	0.548*** (0.006)	0.001
Pr(died by age 60)	0.132*** (0.017)	0.141*** (0.002)	0.573
Share earnings 1. quartile	0.377*** (0.047)	0.290*** (0.005)	0.057
Share earnings 2. quartile	0.274*** (0.026)	0.270*** (0.005)	0.854
Share earnings 3. quartile	0.213*** (0.032)	0.246*** (0.004)	0.286
Share earnings 4. quartile	0.136*** (0.034)	0.193*** (0.006)	0.060
Share income 1. quartile	0.357*** (0.047)	0.287*** (0.006)	0.123
Share income 2. quartile	0.268*** (0.030)	0.269*** (0.005)	0.988
Share income 3. quartile	0.223*** (0.025)	0.243*** (0.004)	0.374
Share income 4. quartile	0.152*** (0.033)	0.199*** (0.005)	0.104

Notes: The table reports the complier characteristics (β -estimates of equation (13)) for a benefit change and compares the marginal entrant to the average entrant. Earnings and income are measured in $t - 3$. The last column reports the p-value for the null hypothesis that marginal entrant characteristics are identical to characteristics of the average entrant. Levels of significance: *10%, **5%, and ***1%.

C Additional Results for Wage Shocks

Figure C.1: 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 Canadian Association of Petroleum Producers (<https://www.capp.ca/en/capp-data-centre/>).

Table C.1: Summary Statistics, Wage Change Samples

	1982–2001	2006–2016
<i>A. Outcomes</i>		
DI benefit receipt (in %)	2.71	2.74
Change in DI benefit receipt (in p.p.)	0.94	0.82
Non-employment (in %)	20.13	18.57
Change in non-employment (in p.p.)	1.28	1.31
Avg. lifetime income	36,600	45,600
	(64,300)	(79,500)
Change in avg. lifetime income	3,800	5,500
	(22,700)	(37,400)
DI benefits of recipients	8,300	8,700
	(4,700)	(4,900)
Earnings	39,400	46,900
	(68,700)	(92,300)
<i>B. Characteristics</i>		
Share Female	0.51	0.52
Share Married	0.63	0.62
Age	39.59	41.55
	(10.20)	(10.82)
No. Kids	0.29	0.35
	(0.64)	(0.70)
No. Observations	9,973,095	8,856,110

Notes: The table reports summary statistics for the wage change samples in 1982-2001 and 2006-2016. Standard deviations of continuous variables in parentheses. All monetary outcomes are expressed in 2019 dollars using the national CPI. Dollar amounts are rounded to \$100 following the vetting guidelines of the research data center.

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

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

Notes: The first two columns report the OLS and 2SLS estimates of equation (11) when we instrument for ΔE_{ict} using the industry shift-share instrument. The last column reports the reduced-form that regresses an outcome variable on the industry shift-share instrument. Coefficient estimates and standard errors for Δ DI enrollment, Δ DI w/o survivor, and Pr(moving) are multiplied by 100 and should be interpreted as percentage points. Standard errors are reported in parentheses and are clustered at the Census Division (CD). The sample includes 15-59 year-old individuals between 1982 and 2016. Levels of significance: *10%, **5%, and ***1%.

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
Obs.	18,829,205	18,829,205	18,829,205
<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.	8,953,440	8,953,440	8,953,440

Notes: The table reports the first-stage and 2SLS estimates of equation (11) for different earnings and income measures when we instrument for ΔE_{ict} using predicted oil employment (panel A) or the oil price (panel B). Coefficient estimates and standard errors for Δ DI enrollment are multiplied by 100 and should be interpreted as percentage points. Standard errors are reported in parentheses and are clustered at the Census Division (CD). The sample includes 15-59 year-old individuals between 1982 and 2016. Levels of significance: *10%, **5%, and ***1%.

Table C.4: The Impact of Lifetime Income on Disability Insurance Enrollment, Robustness

	OLS	Industry share	Oil employment	Oil price
<i>A. Years 1982-2001</i>				
Coefficient enrollment	-0.016*** (0.004)	-0.082*** (0.024)	-0.084*** (0.023)	-0.052*** (0.010)
Effective F-statistic		46.5	14.9	17.4
Obs.	9,973,095	9,973,095	9,973,095	3,386,005
<i>B. Oil and gas provinces</i>				
Coefficient enrollment	-0.013*** (0.005)	-0.035** (0.016)	-0.031** (0.016)	-0.043*** (0.007)
Effective F-statistic		23.0	192.8	34.0
Obs.	9,412,320	9,412,320	9,412,320	4,371,960

Notes: The table reports the OLS and 2SLS estimates of equation (11), instrumenting ΔE_{ict} (lifetime income) using predicted employment based on industry employment shares, oil employment shares, or oil prices. Coefficient estimates and standard errors are multiplied by 100 and should be interpreted as percentage points. Standard errors are reported in parentheses and are clustered at the Census Division (CD). The sample includes 15-59 year-old individuals between 1982 and 2001 (Panel A) and 15-59 year-old individuals between 1982 and 2016 who reside in oil and gas provinces (Alberta, British Columbia, Saskatchewan, and Newfoundland and Labrador) or Quebec (Panel B). Levels of significance: *10%, **5%, and ***1%.

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

	Industry shares			Oil shares
	Technical services	Farming	Transportation	
Age 15-19	0.062 (0.038)	-0.350 (0.425)	-0.092 (0.064)	-0.160 (0.245)
Age 20-24	-0.009 (0.046)	-0.113 (0.373)	-0.055 (0.056)	-0.131 (0.254)
Age 25-29	0.002 (0.051)	0.181 (0.351)	-0.217*** (0.069)	-0.084 (0.248)
Age 30-34	-0.107*** (0.035)	-0.195 (0.340)	-0.010 (0.057)	-0.253 (0.231)
Age 35-39	0.054 (0.040)	-1.559*** (0.446)	0.071 (0.077)	-0.209 (0.259)
Age 40-44	0.067 (0.096)	0.381 (0.814)	-0.171 (0.138)	-0.228 (0.236)
Age 45-49	0.003 (0.076)	-1.501* (0.831)	-0.040 (0.097)	-0.242 (0.257)
Age 50-54	-0.021 (0.073)	0.283 (0.711)	-0.063 (0.132)	-0.312 (0.251)
Age 55-59	0.017 (0.104)	1.292** (0.597)	-0.017 (0.112)	-0.166 (0.250)
Pr(died by age 60)	0.463 (1.618)	-17.488 (12.056)	-1.597 (1.940)	0.183 (1.577)
Share earnings 2. quartile	-0.046 (0.047)	0.620 (0.504)	0.039 (0.086)	0.012 (0.016)
Share earnings 3. quartile	-0.001 (0.069)	1.461** (0.583)	0.052 (0.130)	1.001 (0.856)
Share earnings 4. quartile	0.109 (0.100)	-0.680 (0.604)	-0.007 (0.117)	0.581 (1.228)
Share income 2. quartile	0.065 (0.048)	-0.258 (0.585)	-0.055 (0.097)	0.053 (0.328)
Share income 3. quartile	0.030 (0.072)	-1.514** (0.646)	0.007 (0.133)	-0.458 (1.020)
Share income 4. quartile	-0.038 (0.106)	0.687 (0.665)	0.022 (0.133)	-0.422 (1.173)
Family size	-0.014*** (0.003)	-0.008 (0.024)	-0.001 (0.004)	-0.009** (0.004)
Share married	-0.015 (0.018)	0.676*** (0.133)	0.021 (0.030)	0.047*** (0.017)
Share female	0.095* (0.055)	-0.376 (0.319)	0.119** (0.058)	0.001 (0.045)
R^2	0.89	0.73	0.41	0.25

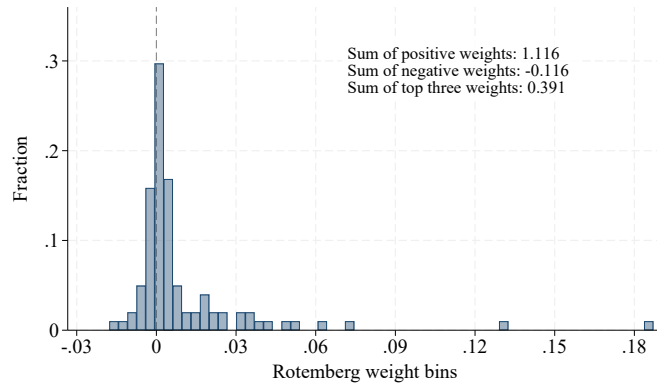
Notes: Each column reports the estimates of a single regression of the initial period the complier characteristics (β -estimates of equation (13)) for benefit and wage changes. Levels of significance: *10%, **5%, and ***1%. Notes: Each column reports results of a single regression of a 1981 industry share on 1982 characteristics. Standard errors in parentheses.

Table C.6: Alternative Estimators and Overidentification Tests

	Coefficient	Over ID test
2SLS (Bartik instrument)	-0.096*** (0.018)	
2SLS (many instruments)	-0.088*** (0.012)	1,020.99 [0.000]
MB2SLS	-0.092*** (0.013)	
LIML	-0.114*** (0.020)	2,491.45 [0.000]

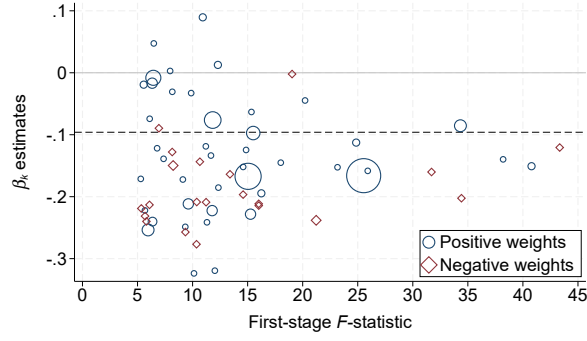
Notes: The table reports the estimates of equation (11) for different estimators, instrumenting for ΔE_{ict} (lifetime income) using predicted employment based on industry employment shares. The dependent variable is the change in DI receipt. The 2SLS estimate with the Bartik instrument reports the main individual-level estimate. The many-instrument 2SLS, MB2SLS, and LIML rows use each industry share (times time period) separately as instruments. Coefficient estimates and standard errors are multiplied by 100 and should be interpreted as percentage points. Standard errors are reported in parentheses and are clustered at the Census Division (CD). P-values for the overidentification tests are reported in brackets. The sample includes 15-59 year-old individuals between 1982 and 2016. Levels of significance: *10%, **5%, and ***1%.

Figure C.2: Distribution of Rotemberg Weights



Notes: This figure shows the full distribution of Rotemberg weights by industry, with weights aggregated across years following Goldsmith-Pinkham, Sorkin and Swift (2020).

Figure C.3: Heterogeneity of β_k -estimates



Notes: The figure relates each industry-share instrument's estimated $\hat{\beta}_k$ to its first-stage strength and Rotemberg weight. Each point represents one industry-share instrument, with $\hat{\beta}_k$ plotted on the y-axis and the corresponding first-stage F-statistic on the x-axis. Marker size is proportional to the absolute value of the Rotemberg weight. Circles indicate positive Rotemberg weights, while diamonds indicate negative weights. The horizontal dashed line shows the overall $\hat{\beta}_k$ from the 2SLS (Bartik) estimate reported in column 2 of Table 2. Instruments with first-stage F-statistics below 5 are excluded.

Table C.7: Comparison Marginal vs. Average Entrant, Wage Change (Industry Share Instrument)

	Marginal entrant	Average entrant	P-value
Age	53.12*** (2.15)	45.69*** (0.16)	0.001
Family size	3.14*** (0.17)	2.69*** (0.02)	0.011
Share female	0.409*** (0.054)	0.585*** (0.003)	0.001
Pr(died by age 60)	0.119*** (0.019)	0.092*** (0.001)	0.143
Share earnings 1. quartile	0.441*** (0.058)	0.252*** (0.003)	0.001
Share earnings 2. quartile	0.256*** (0.050)	0.286*** (0.005)	0.575
Share earnings 3. quartile	0.238*** (0.034)	0.262*** (0.002)	0.482
Share earnings 4. quartile	0.065 (0.063)	0.200*** (0.005)	0.040
Share income 1. quartile	0.345*** (0.040)	0.212*** (0.003)	0.001
Share income 2. quartile	0.292*** (0.053)	0.308*** (0.003)	0.770
Share income 3. quartile	0.264*** (0.030)	0.271*** (0.002)	0.823
Share income 4. quartile	0.099 (0.062)	0.210*** (0.004)	0.086

Notes: The table reports the complier characteristics (β -estimates of equation (13)) for wage changes and compares the marginal entrant to the average entrant. Earnings and income are measured in $t - 3$. The change in DI receipt on the right-hand side of equation (13) is instrumented with the industry shift share instrument. The last column reports the p-value for the null hypothesis that marginal entrant characteristics are identical to characteristics of the average entrant. Levels of significance: *10%, **5%, and ***1%.

Table C.8: Comparison of Complier Characteristics, Wage Change (Oil Instrument)

	Benefit change		Wage change		Equality test (P-value)
	Coeff.	Std. err.	Coeff.	Std. err.	
Age	50.57***	0.81	49.43***	0.93	0.356
Family size	3.60***	0.26	2.82***	0.09	0.004
Share female	0.426***	0.038	0.545***	0.014	0.003
Pr(died by age 60)	0.132***	0.017	0.078***	0.006	0.003
Share earnings 1. quartile	0.377***	0.047	0.376***	0.020	0.983
Share earnings 2. quartile	0.274***	0.026	0.189***	0.034	0.044
Share earnings 3. quartile	0.213***	0.032	0.285***	0.016	0.043
Share earnings 4. quartile	0.136***	0.034	0.150***	0.024	0.730
Share income 1. quartile	0.357***	0.047	0.229***	0.015	0.010
Share income 2. quartile	0.268***	0.030	0.302***	0.028	0.417
Share income 3. quartile	0.223***	0.025	0.288***	0.018	0.036
Share income 4. quartile	0.152***	0.033	0.181***	0.025	0.475

Notes: The table reports the complier characteristics (β -estimates of equation (13)) for benefit and wage changes. The change in DI receipt on the right-hand side of equation (13) is instrumented with $RoC_p \cdot A87_t$ for the benefit change and oil employment instrument for the wage change. Earnings and income are measured in $t - 3$. The last column reports the p-value for the null hypothesis that marginal entrant characteristics are identical for benefit and wage changes. Levels of significance: *10%, **5%, and ***1%.

Table C.9: Comparison Marginal vs. Average Entrant, Wage Change (Oil Instrument)

	Marginal entrant	Average entrant	P-value
Age	49.43*** (0.93)	45.69*** (0.16)	0.000
Family size	2.82*** (0.09)	2.69*** (0.02)	0.110
Share female	0.545*** (0.014)	0.585*** (0.003)	0.002
Pr(died by age 60)	0.078*** (0.006)	0.092*** (0.001)	0.027
Share earnings 1. quartile	0.376*** (0.020)	0.252*** (0.003)	0.000
Share earnings 2. quartile	0.189*** (0.034)	0.286*** (0.005)	0.002
Share earnings 3. quartile	0.285*** (0.016)	0.262*** (0.002)	0.114
Share earnings 4. quartile	0.150*** (0.024)	0.200*** (0.005)	0.028
Share income 1. quartile	0.229*** (0.016)	0.212*** (0.003)	0.195
Share income 2. quartile	0.302*** (0.028)	0.308*** (0.003)	0.840
Share income 3. quartile	0.288*** (0.018)	0.271*** (0.002)	0.335
Share income 4. quartile	0.181*** (0.025)	0.210*** (0.004)	0.223

Notes: The table reports the complier characteristics (β -estimates of equation (13)) for wage changes and compares the marginal entrant to the average entrant. Earnings and income are measured in $t - 3$. The change in DI receipt on the right-hand side of equation (13) is instrumented with the oil employment instrument. The last column reports the p-value for the null hypothesis that marginal entrant characteristics are identical to characteristics of the average entrant. Levels of significance: *10%, **5%, and ***1%.

Table C.10: Comparison of Complier Characteristics, Wage Change (Industry Share Instrument, 1982–2001)

	Benefit change		Wage change		Equality test (P-value)
	Coeff.	Std. err.	Coeff.	Std. err.	
Age	50.57***	0.81	48.76***	2.10	0.422
Family size	3.60***	0.26	3.33***	0.26	0.454
Share female	0.426***	0.038	0.374***	0.051	0.422
Pr(died by age 60)	0.132***	0.017	0.119***	0.018	0.604
Share earnings 1. quartile	0.377***	0.047	0.449***	0.059	0.338
Share earnings 2. quartile	0.274***	0.026	0.283***	0.073	0.913
Share earnings 3. quartile	0.213***	0.032	0.157***	0.035	0.234
Share earnings 4. quartile	0.136***	0.034	0.112	0.092	0.805
Share income 1. quartile	0.357***	0.047	0.431***	0.052	0.294
Share income 2. quartile	0.268***	0.030	0.234***	0.089	0.719
Share income 3. quartile	0.223***	0.025	0.186***	0.031	0.357
Share income 4. quartile	0.152***	0.033	0.149	0.092	0.978

Notes: The table reports the complier characteristics (β -estimates of equation (13)) for benefit and wage changes for the time period 1982–2001. The change in DI receipt on the right-hand side of equation (13) is instrumented with $RoC_p \cdot A87_t$ for the benefit change and the industry shift-share instrument for the wage change. Earnings and income are measured in $t - 3$. The last column reports the p-value for the null hypothesis that marginal entrant characteristics are identical for benefit and wage changes. Levels of significance: *10%, **5%, and ***1%.

Table C.11: Comparison of Complier Characteristics, Wage Change (Oil Instrument, 1982–2001)

	Benefit change		Wage change		Equality test (P-value)
	Coeff.	Std. err.	Coeff.	Std. err.	
Age	50.57***	0.81	50.62***	0.83	0.967
Family size	3.60***	0.26	2.86***	0.11	0.008
Share female	0.426***	0.038	0.522***	0.022	0.029
Pr(died by age 60)	0.132***	0.017	0.088***	0.011	0.027
Share earnings 1. quartile	0.377***	0.047	0.416***	0.017	0.437
Share earnings 2. quartile	0.274***	0.026	0.166***	0.039	0.022
Share earnings 3. quartile	0.213***	0.032	0.255***	0.016	0.242
Share earnings 4. quartile	0.136***	0.034	0.163***	0.026	0.533
Share income 1. quartile	0.357***	0.047	0.240***	0.018	0.020
Share income 2. quartile	0.268***	0.030	0.321***	0.036	0.260
Share income 3. quartile	0.223***	0.025	0.240***	0.029	0.661
Share income 4. quartile	0.152***	0.033	0.199***	0.028	0.271

Notes: The table reports the complier characteristics (β -estimates of equation (13)) for benefit and wage changes for the time period 1982–2001. The change in DI receipt on the right-hand side of equation (13) is instrumented with $RoC_p \cdot A87_t$ for the benefit change and the oil employment instrument for the wage change. Earnings and income are measured in $t - 3$. The last column reports the p-value for the null hypothesis that marginal entrant characteristics are identical for benefit and wage changes. Levels of significance: *10%, **5%, and ***1%.

D Additional Results for Welfare Analysis

Table D.12: Welfare Analysis, Years 1982-2001

<i>A. Δ DI take-up 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)	-0.082*** (0.024)	-0.084*** (0.023)	-0.052*** (0.010)
<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)	2.542*** (0.810)	2.485*** (0.733)	4.020*** (0.881)
P-value: estimate = 1	0.000	0.057	0.043	0.001
P-value: multiplier = ins. value		0.227	0.207	0.004

Notes: The table combines the estimates from column 1 of Table 1 and Panel A of Table C.4 to construct the multiplier and insurance value of higher DI benefits. Levels of significance: *10%, **5%, and ***1%.

Table D.13: Welfare Analysis, Oil and Gas Provinces

<i>A. ΔDI take-up per \$1,000</i>				
	DI benefits ($\partial DI/\partial b$)	Lifetime income ($\partial DI/\partial w$)		
		Industry share	Oil employment	Oil price
Coeff. estimate	0.160*** (0.022)	-0.035** (0.016)	-0.031** (0.016)	-0.043*** (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	1.515*** (0.071)	4.531** (2.140)	5.073* (2.638)	3.720*** (0.800)
P-value: estimate = 1	0.000	0.099	0.123	0.001
P-value: multiplier = ins. value		0.155	0.174	0.004

Notes: The table combines the estimates from column 2 of Table B.2 and Panel B of Table C.4 to construct the multiplier and insurance value of higher DI benefits. Levels of significance: *10%, **5%, and ***1%.