Occupational Hazards and Social Disability Insurance

Amanda Michaud[a,*], David Wiczer[b]

[a]*Indiana University*
[b]*Federal Reserve Bank of St Louis*

Abstract

Using retrospective data, we introduce evidence that occupational exposure significantly affects disability risk. Incorporating this into a general equilibrium model, social disability insurance (SDI) affects welfare through (i) the classic, risk-sharing channel and (ii) a new channel of occupational reallocation. Both channels can increase welfare, but at the optimal SDI they are at odds. Welfare gains from additional risk-sharing are reduced by overly incentivizing workers to choose risky occupations. In a calibration, optimal SDI increases welfare by 2.6% relative to actuarially fair insurance, mostly due to risk sharing.

JEL CODES: E62, I13

Keywords: Disability Insurance, Occupational Choice, Optimal Policy

1. Introduction

Social insurance against disability risk is a substantial and growing part of Federal expenditures in the United States. Social Security disability insurance (SSDI) provided benefits to 4.7% of working-age adults and cost about $0.2 trillion (1.7% of GDP) in 2012. For comparison, the next largest non-employment social insurance program, unemployment insurance (UI), benefited 2.5% of the working-age population at a cost of $78.9 billion in 2012.[1] Evaluating the social benefits of this program, however, one must consider the nature of disability risk. We provide evidence that lifetime occupational choices are an important component of individuals' physical disability risk. The welfare effects of social disability insurance is then affected by the ability of workers to choose occupations considering these risks.

Evaluation of social disability insurance (SDI) programs often focuses on the trade-off between insurance and incentives. From this view, disability insurance benefits society by providing a valuable transfer to workers who suffer an adverse shock (e.g., Gruber (1997)). However, insurance is costly because it provides incentives to misrepresent one's disability status. This paper considers a different incentive created by disability insurance. When the nature of work affects disability risk, SDI alters the composition of occupations in the US economy by incentivizing workers to chose occupations with higher risk.

We evaluate SDI within a theoretical model of occupation choice and occupation-specific disability risk, highlighting the importance of general equilibrium effects. We show that, in a world with no insurance (incomplete markets), fewer workers choose occupations with high disability risk than the social planner would allocate. In our quantitative analysis, we show actuarially fair insurance *also* places fewer workers in risky occupations than the social planner would. This occurs because workers in high-risk jobs demand a sufficiently high wage to self-insure against future inability to work. In the presence of a decreasing marginal product of additional workers, these high wages are maintained by shifting workers to less-risky occupations. If occupations are not perfect substitutes, this imbalance makes production inefficient. Equilibrium prices cannot achieve efficiency because self-insurance is too expensive.

The introduction of SDI improves welfare through two channels. The first is the classical

---

[1]UI beneficiaries ranged from 1.7-4.5% of the covered workforce from 2000 to 2013. (from DOL ET Financial Data Handbook 394 Report (2013) and Administration (2013).

risk sharing channel: social insurance improves welfare by helping workers in risky occupations to smooth consumption. The second is a novel channel of occupational reallocation: SDI encourages more workers to choose risky occupations. As a result, the allocation of workers across occupations becomes more efficient, output increases, and consumption becomes cheaper. These benefits are larger than the cost to fund the scheme because social insurance is more efficient than self-insurance. With self-insurance some workers who accumulate savings never become disabled. With SDI, funds need to be provided only for workers who actually become disabled. Improved productive efficiency results in welfare gains for workers even with zero disability risk and many of those who never experience disability.

The two channels, risk-sharing and the reallocation of workers, increase welfare upon the marginal introduction of social insurance, but they are at odds at the welfare-maximizing level. Although reallocation towards risky occupations initially improves productive efficiency, and output maximizing level exists and too much reallocation can compromise it. At the point at which increases in benefits reduce output, there are still welfare gains from risk-sharing. This result is actually general; welfare is maximized by an SDI program that induces workers to choose risky occupations beyond the output maximizing level.

We show that our results are qualitatively robust to ex ante heterogeneity in preferences that induce sorting and to private information over disability status, though it introduces bounds on the benefit. Further, private information over disability status, which limits the dimensions on which insurance can contract, is a reason for why private insurance cannot achieve the same welfare gains as social insurance. If neither public nor private contracts can condition on workers' risk-level, i.e. contracts are not occupation-specific, high-risk workers over subscribe to private insurance contracts, whereas the public system can dictate equal insurance coverage. This adverse selection leads to a market failure typical of insurance contracts with private information.[2] Social insurance with compulsory enrollment is not subject to adverse selection and always generates welfare gains, as we describe.

Our theoretical work is motivated by empirical work that shows heterogeneity in disability incidence. Using data from the University of Michigan's Health and Retirement Survey (HRS), we find a natural grouping between low- and high-risk occupations, the latter have

---

[2]Market failure here is Rothschild Stiglitz-like, in that if an insurance contract were offered, it would attract high-risk workers and earn a negative profit. Increasing the price only makes adverse selection worse.

about twice the disability rate as the low-risk group. While this is suggestive, occupational choice is endogenous and potentially influenced by unobservable factors. We therefore propose an instrument scheme using O*NET measures of physical and non-physical occupational requirements. Physical requirements have health repercussions, but the estimated effect of these requirements may be influenced by sorting along just these physical requirements. Therefore, we instrument physical requirements by the non-physical requirements in the O*NET. Our instrument works because occupations bundle requirements: Even though workers may sort along the physical requirements and sully the estimate of their effect, we can instrument with the non-physical requirements which are found bundled together. This instrumenting scheme is fruitful: it uncovers small but significant bias in the estimate assuming exogeneity and, via an over-identification J-test, cannot be rejected as valid.

We use these facts, along with the US SSDI system, to calibrate our model of occupation risk. We find the optimal SDI program costs 1.45% of GDP and provides a replacement rate of 59% to disabled workers. It provides welfare gains equivalent to a 2.6% increase in consumption in a world with actuarially fair insurance alone. Relative to this optimal program, the current US system captures 92% of the potential gains. We conclude from these findings that there is a quantitatively important role for SDI beyond the insurance that is provided by private markets.

The economic mechanism of this paper is most related to that of Acemoglu and Shimer (1999). They show unemployment insurance can raise output by inducing workers to search for higher productivity jobs which are rarer, and therefore, more risky to pursue. In this paper, SDI also increases output by inducing workers to take on more risk in their job search; specifically choosing occupations with greater disability risk. This common mechanism provides a unique motivation for social insurance beyond that offered by private markets. Social insurance is not simply a transfer to those experiencing bad luck. Instead, improvements in productive efficiency increase the welfare of all individuals, even those who face little or no risk.

Several papers have discussed similar topics but with elements missing from our analysis. Golosov and Tsyvinski (2006) also normatively consider optimal DI, but from a mechanism design approach to prevent workers from misreporting their disability status. We show workers would truthfully report their status in our baseline SDI scheme, but it is because our policy tool is already constrained from providing full-insurance. On the other side of the

related literature, many authors have discussed how claiming disability insurance is affected by employment prospects with more detailed labor markets (e.g. Rust and Phelan, 1997; Duggan and Autor, 2006; Kitao, 2014). We omit these and other interesting factors from our study, to help keep our analysis of the main mechanisms clear. Schulhofer-Wohl (2011) also considers workers who choose jobs with different levels of risk and focuses on underlying heterogeneity in preferences. He shows that this generalization can reduce the welfare costs associated with incomplete insurance. In an extension, we incorporate preference heterogeneity and show our SDI scheme still generates Pareto-improving welfare gains, although those gains may now be unequal. Quite consistent with the results of Schulhofer-Wohl (2011) Workers in the most risky occupations may gain the least from insurance because the most impatient agents select into those occupations.

We begin by motivating our normative work by estimating the disability risk associated with various occupations. A few papers have addressed a similar empirical task (e.g. Fletcher et al., 2009; Morefield et al., 2011). These papers both connect physically demanding occupations to health problems later in life. Ravesteijn et al. (2013) use German data and, again, link occupational physical demands to health deteriorations. They use a dynamic fixed effects model to control the individual effects affecting both occupational choice and health outcomes. We introduce a set of instrumental variable techniques to specifically address this problem of endogeneity between occupational choice and potential heterogeneity in risk sensitivity.

2. Data on occupations and disability

In this section, we present data regarding the connection between an individual's occupation and disability risk. First, we construct a measure of lifetime exposure to an occupation using the University of Michigan's Health and Retirement Study (HRS). The survey respondents provide detailed information on their health conditions, from which we infer whether they are "disabled." We find that occupations have quite disparate disability rates, much of which can be attributed to their occupations' effects themselves.

2.1. Constructing the dataset

We begin with RAND's distribution of the HRS. The HRS surveys households whose head is older than 50 years of age and also includes data on their "spouses," potentially an unmar-

ried cohabitant. The panel is collected biannually from 1992 to 2010. The individual listed as a spouse may change through the survey waves and we drop these observations. After excluding observations for non-responses, we have between 16,128 and 21,623 observations per wave with a total of 184,541.

We use one of two measures of disability. Our baseline uses the more restrictive measure, which looks at whether they report limitations to activities of daily living (ADLs). The broader of the two uses the survey's direct question, whether a health problem limits one's ability to do paid work. Results are generally very similar between the two and we replicate all of our estimation with the latter measure in Appendix Section A.3. For each measure, we count only health problems if they occur during working life, before 65.

In each wave, individuals report difficulties they have across many ADLs. These include walking across a room, getting dressed, bathing, and getting in and out of bed. We record an indicator once a worker begins to continuously report a difficulty with any ADL.[3]

We will associate an individual with his or her longest-held occupation, reported retrospectively. If we used the current occupation, we would mistakenly consider an occupation dangerous if workers in poor health switched into that occupation later in life. Because "longest-held" is repeatedly asked each wave, we take the occupation with the longest tenure.[4] The longest-held occupation has a median tenure of 19 years and the bottom quartile is 11 years.

For each occupation, we merge in Knowledge, Skills and Abilities descriptors from O*NET using the analyst database (version 4.0). We then split these descriptors between the physical demands of an occupation and the rest, 19 of the former and 101 of the latter. From the physical descriptors, we measure the occupation's physical demands using the first principal component. For the 101 other descriptors, we compute 3 principal components, about 70% of the variation. Finally, we de-mean and standardize each of the components.

---

[3]We also experimented with a metric over the number and severity of the difficulties. Results are all robust to this definition and are available upon request.

[4]The longest-held occupation may change from wave to wave because the individual's tenure in the current occupation overtakes the prior longest-held occupation or because of coding error. The latter occurs 4% of the time in the first wave, but then less than 1% of the time. We overwrite these coding errors.

2.2. The distribution of disability across occupations

Occupations differ greatly in their average disability rates. Particularly, there is quite a long tail of high-risk occupations while most of the population works in occupations clustered at fairly low levels of risk. The distinction between high-risk and low-risk occupations is great enough that the distribution appears bimodal.

Figure 1 shows the distribution across occupations of the incidence of a disability, as measured by having a difficulty with an ADL. In this kernel density estimate, each data point is an occupation, and within that occupation, we construct the average rate of ADL difficulties. We then weight the occupations by their population using the appropriate weights provided by the HRS. Along the horizontal axis, we plot dots representing the position of occupations.
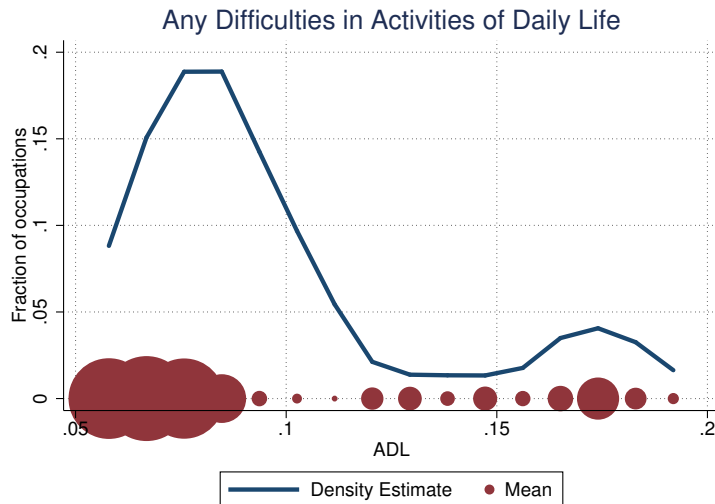
Figure 1: The density across occupations of the incidence of difficulties with ADLs

Note the large clustering of occupations at the low end of the distribution and the long tail of occupations with more than twice the average rate of ADL difficulties. To summarize these two groups, we estimate a mixture of two normal distributions to classify occupations as "risky" or "safe." From our primary measure, whether a worker has any limitations to ADLs, the mean disability rate in the high risk group is about 16%, twice that of the the low-risk group. About 17% of the population worked in this high risk group.

The largest high-risk occupations are machine and transport operators. Construction and extraction also represent a large portion of the high-risk occupations, though these

occupations are under-represented compared to a cross-section because many work in these occupations temporarily and are not their longest-held. The large, low-risk occupations are professional and management occupations, administrative support and sales.

2.3. Extracting the occupation effect

How much of the difference in disability outcomes across occupations can be ascribed to the occupations themselves? Potentially, there are systematic differences between those who choose what appear to be "risky" occupations and those who choose safe occupations. We attempt to find an effect of occupation on disability in two ways. We use an Oaxaca-Blinder decomposition to extract the residual effect of occupation on disability from observable differences between workers in different occupations. However, this decomposition is problematic because the grouping is endogenous, i.e. people choose their occupations. We address this issue by using instrumental variable methods: We use O*NET measures of physical demands to summarize the occupation's effect on health and then instrument these physical demands by non-physical demands that tend to be bundled within occupations.

In our first method to extract the occupation effect on disability, we segment occupations into low- and high-risk based on the mixture we estimated on Figure 1, assigning each occupation to the distribution with highest density at its level of risk. Then we use a Oaxaca-Blinder decomposition to isolate the difference due to observable characteristics and the residual differences in disability rates across occupation groups.

On the left-hand side is whether the individual ever reported disability and on the right-hand side our regressors are a cubic for potential experience, body mass index (BMI), time, dummies for education level, gender, marital status, race, and tobacco use. For workers who report a disability, we take a snapshot at the wave in which they become disabled. We use the mean value across waves for workers who never report a disability. The results of the decomposition are summarized in Table 1.

Column (1) of Table 1 uses only one observation per individual, with covariates averaged over the panel; column (2) pools all of the data with multiple observations per individual. In our baseline case, Column (1), the occupation effect accounts for 42% of the difference in risk. The finding is robust to using health limitations (Appendix A.3), where the riskier occupations have much higher rates of limitations, 35% of which is due to the residual effect of the occupations themselves. From this decomposition, we see that much of the difference

	(1)	(2)
Safe	0.189	0.106
Risky	0.264	0.157
Difference	-0.076	-0.051
Observables	-0.044	-0.030
% Difference	57.9	58.8
Occupation	-0.031	-0.021
% Difference	42.1	41.2
N	20,328	127,298

Table 1: The decomposition of the occupation-group effect on disability. Occupations are split between low- and high-risk and the regressors are a cubic for potential experience, body mass index (BMI), time, dummies for education level, gender, marital status, race, and tobacco use. Column (1) uses only one observation per individual and (2) pools all of the data.

across occupation groups cannot be explained by differences in observables across the two groups.

While this decomposition is instructive, the fact that so many observable characteristics differ between workers in risky and safe occupations implies there also may be sorting on unobservable characteristics. If, for example, workers who are more tolerant of disability, and therefore engage in more risky activities outside of work, sort disproportionately into high-risk occupations we would over-state the occupation effect. On the other hand, if more physically robust workers sort into physically-demanding occupations, this would attenuate occupations' effect. We now use instrumental variables to try to address this.

We begin by mapping occupations' health risk into the physical demands characterized by the O*NET, the construction of which we described in subsection 2.1. To relate back to our two occupation groups, risky occupations have a mean physical requirement of 1.05 and safe occupations have a mean of -0.65.[5] Using a direct measure of physical requirement instead

---

[5]To interpret these numbers, recall, we standardized this measure to mean 0 and standard deviation 1

of the the identify of the occupation solves one problem: if there is a characteristic of risky occupations correlated with disability outcomes and more common among risky occupations. For example, risky occupations may be over-represented in declining industries where poor worker prospects diminish worker's labor-force attachment and increase their likelihood of reporting a disability. Characterizing the occupation only by its physical requirement cuts away these potential other characteristics. However, this mapping does not address sorting on the desired level of health risk itself, the fundamental endogeneity problem we face.

To address this, we create a set of instrumental variables from the non-physical requirements of occupations. Because occupations bundle many requirements, and there are patterns in these bundles, we can use the non-physical requirements to predict the physical requirements of an occupation. We can expect these requirements to be valid instruments because these non-physical requirements should be uncorrelated with one's propensity for disability. To make clear our strategy, for an individual i in occupation j, we model the probability of ever becoming disabled before age 65 as

$$\Pr[\text{Disabled}_i] = f(\gamma H_j + x_i'\beta + \nu_i) \tag{2.1}$$

where x_i is the same vector of observable characteristics as our previous specification and f is the link function. H_j is the first principal component of the physical descriptors associated with occupation j and ν_i is an unobservable individual trait, physical "robustness." Robustness may affect both the probability of disability and the level of physical requirements one chooses in an occupation. We therefore treat H_j as endogenous and instrument it with the first 3 principal components of the non-physical requirements of the occupation.[6] The idea behind our instrument is that workers may choose their occupation considering ν_i, e.g. particularly athletic individuals may chose more physically intense jobs given they know they are less sensitive to these demands, but they do not chose other non-physical occupational characteristics based on ν_i. Occupations, however, follow certain patterns in how

---

[6]The number of components we use as instruments is not completely innocuous. Doubling the number of components, they predict the physical requirements near perfectly and invalidate the instruments. Our results are qualitatively the same if we increase the number of principal components to 4 or reduce to 2, though reducing the number of instruments diminishes the power of our tests of over-identification and therefore our ability to test the instruments' validity.

they bundle requirements, so the non-physical requirements can predict well the physical requirements.

The marginal effects from the estimation of Equation 2.1 are in Table 2. Column (1) displays the Probit estimates and Column (2) displays the linear probability model, both using our IV scheme. Columns (3) and (4) treat physical requirements as exogenous in a probit and LPM, respectively. The first stage is highly predictive, the p-value associated with its F-statistic is 0.000 and, as would be expected, the first principal component of non-physical requirements is strongly negatively related to our measure of physical requirements. Estimates are available in the Appendix Section A.1. For robustness, in Appendix Section A.2 we also include estimates including industry dummies and clustering standard errors on industry, which may be important to control for economic factors affecting labor-force attachment and disability reporting. This changes the results quantitatively very little. Appendix section A.3 includes estimates using our alternative disability definition, health limitations to work. Again, the results are qualitatively robust.

The principal result from these estimates is in the first row and is quite consistent across estimation methods. In our baseline specification, a standard deviation change in physical requirements increases disability risk by 3.4 percentage points. Applying this coefficient to our two groups of occupations: the higher physical requirement of risky occupations implies about a 5.7 percentage point higher rate of ADL difficulty than other occupations. Notice that the instruments increase the margin by a small, but statistically significant amount, which suggests that people sort into more physically demanding occupations if they have a lower disability risk. The instruments also pass our over-identification test. We list Hansen's J-statistic for the linear IV in Column (2) and Amemiya-Lee-Newey for the Probit in Column (1). With both estimates, we clearly cannot reject the validity of these instruments.

	(1)	(2)	(3)	(4)
O*NET Phys Reqs	0.0338 **	0.0365 **	0.0315 **	0.0344 **
	(0.0048)	(0.0050)	(0.0045)	(0.0048)
Potential Experience	0.0327 **	0.0323 **	0.0328 **	0.0323 **
	(0.0007)	(0.0008)	(0.0007)	(0.0008)
BMI	0.0122 **	0.0148 **	0.0122 **	0.0148 **
	(0.0007)	(0.0009)	(0.0007)	(0.0009)
Woman	0.0272 **	0.0457 **	0.0258 **	0.0447 **
	(0.0085)	(0.0082)	(0.0085)	(0.0082)
Self-employed	-0.0222 *	-0.0141	-0.0224 *	-0.0143
	(0.0099)	(0.0088)	(0.0099)	(0.0088)
College	0.1436 **	0.1543 **	0.1425 **	0.1532 **
	(0.0100)	(0.0101)	(0.0099)	(0.0101)
No HS	-0.0927 **	-0.0827 **	-0.0917 **	-0.0818 **
	(0.0111)	(0.0114)	(0.0111)	(0.0114)
Not married	0.0865 **	0.0959 **	0.0867 **	0.0960 **
	(0.0106)	(0.0130)	(0.0106)	(0.0130)
Not white	0.0310 **	0.0413 **	0.0315 **	0.0418 **
	(0.0103)	(0.0121)	(0.0102)	(0.0121)
Smoker	0.0210 *	0.0246 **	0.0214 *	0.0248 **
	(0.0094)	(0.0095)	(0.0095)	(0.0095)
Observations	14763	14763	14763	14763
J statistic	0.4749	0.9753		
p-value, J	0.7886	0.6141		

Robust standard errors in parentheses

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 2: The effect of an occupation's physical requirements on working-life disability. Columns (1) and (2) use our instrumental variable schemes and Columns (3) and (4) treat physical requirements as exogenous. (1) and (3) are marginal effects at the mean from probit models. (2) and (4) are coefficients of linear probability models.

3. A Model Economy with Occupation-Specific Disability Risk

To prove our theoretical results, we use a two-period overlapping generations model where disability status is a publicly-observed, discrete event. Workers who receive a disability shock cannot work.

Time is discrete. There is a single consumption good produced with labor by workers in a continuum of different occupations $j \in [0, J]$. The production technology exhibits constant elasticity of substitution across occupations with elasticity of substitution $\rho = \frac{1}{1-\gamma} < \infty$. The index of an occupation also defines an occupation-specific probability of disability in the second period: $\theta(j) = j$.

There is a continuum of competitive firms. The representative firm hires labor n_j in occupation specific spot markets to solve the following maximization problem:[7]

$$\max_{\{n(j)\}_{j=0}^J} y - \int_{j=0}^{J} w(j)n(j)dj \text{ s.t.} \tag{3.1}$$

$$y = \left(\int_{j=0}^{J} n(j)^\gamma \right)^{1/\gamma} \tag{3.2}$$

Each period a unit measure of workers is born. Workers are identical at birth and live for two periods. They have strictly risk-averse, time-separable preferences over consumption in both periods $U(c_1, c_2) = u(c_1) + u(c_2)$. The utility function $u(\cdot)$ is assumed to be strictly increasing, strictly concave, homothetic, and continuously differentiable.

In the first period, workers choose a single occupation, which persists their entire lifetime. They then work, collecting wage earnings net of taxes $(1 - \tau)w(j)$. From these earnings, they decide how much to consume, c_1, and how much to save in a, a storage technology with rate of return 1.

In the second period, disability shocks occur with occupation-specific probability $\theta(j)$. Individuals who receive a disability shock cannot work and report themselves as disabled. Their income is disability benefits with replacement rate b; ie: $bw(j)$. Individuals who do not receive a disability shock will work. Their income is wage earnings net of taxes

---

[7]Because firms hire in spot markets, their problem is static. This means they cannot design complicated, multi-period wage contracts. It also implies firms do not internalize that hiring workers in risky occupations decreases labor available in the next period.

$(1 - \tau)w(j)$. Agents consume whatever income plus savings they have, then die. Subscript 1 denotes period 1 consumption, d is period 2 consumption when a worker is disabled and n is for period 2 when a worker is not disabled. This problem can be represented as $\max_{j \in [0,J]} E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))]$, where we use E_j to show the expectation over risks associated with occupation j, $\theta(j)$.

$$E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))] = \max_{c_1, c_n, c_d, a} u(c_1) + (1 - \theta(j))u(c_n) + \theta(j)u(c_d) \text{ s.t.} \tag{3.3}$$

$$c_1 \leq w(j)(1 - \tau) - a \tag{3.4}$$

$$c_n \leq w(j)(1 - \tau) + a \tag{3.5}$$

$$c_d \leq a + bw(j) \tag{3.6}$$

$$a \geq 0 \tag{3.7}$$

The solution to this problem gives occupation-specific decision rules: $c_1^*(j), c_d^*(j), c_n^*(j), a^*(j)$. Define $\ell^*(j)$ as the measure of workers choosing occupation j .

Definition 3.1 (Competitive equilibrium). An equilibrium consists of allocations $\{c_1^*(j), c_d^*(j), c_n^*(j), a^*(j), \ell^*(j), n^*(j)\}$, government policies $\{\tau, b\}$, and prices $w^*(j)$ for every $j \in [0, J]$ such that (i) given prices and government policies, allocations solve the workers' and firms' problems; (ii) feasibility is satisfied in the labor market: $n^*(j) \leq (2 - \theta(j))\ell^*(j)$; markets clear in (iii) goods; and (iv) government budgets balance period-wise.

Optimality in the labor market requires satisfying two conditions:

- Firms: Wage equals marginal product: $w(j) = \left(\frac{y}{n(j)}\right)^{1-\gamma}$

- Workers: Indifferent between entering any occupation:

$$E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))] = E_k[Uc_1^*(k), c_n^*(k), c_d^*(k))] \ \forall j, k \in [0, J] \ .$$

Goods market clearing requires that

$$y \geq \int_{j=0}^{J} c_1(j) + (1 - \theta(j))c_n(j) + \theta(j)c_d(j)dj$$

Government budget balance requires

$$b^* \leq \frac{\tau y}{\int_j \theta(j)\ell(j)^*w(j)^*dj}$$

3.1. Sources of Inefficiency in the Competitive Equilibrium

Our first proposition shows that the competitive allocation without SDI ($\tau = 0$) is Pareto inefficient. A first-best social planner can choose an alternative feasible allocation that improves welfare for all agents.

Proposition 3.2. *Let $\{c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j), \ell^{cm}(j)\}$ be the first-best planner's allocation solving*[8]

$$\max_{\{c_1(j), c_d(j), c_n(j), \ell(j)\}} \int_j \ell(j) \left(u(c_1(j)) + \theta(j) u(c_d(j)) + (1 - \theta(j)) u(c_n(j)) \right) dj \ \ s.t. \tag{3.8}$$

$$\left(\int_j ((2 - \theta(j)) \ell(j))^\gamma dj \right)^{\frac{1}{\gamma}} \geq \int_j \ell(j) \left(c_1(j) + \theta(j) c_d(j) + (1 - \theta(j)) c_n(j) \right) dj \tag{3.9}$$

$$1 \geq \int_j \ell(j) dj \tag{3.10}$$

Then, $\{c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j), \ell^{cm}(j)\}$ strictly Pareto dominates $\{c_1^*(j), c_d^*(j), c_n^*(j), \ell^*(j)\}$.

Proof. See Appendix B.4 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

The first-best planner's allocation provides welfare gains through two channels: (i) reallocation of consumption across workers (risk-sharing) and (ii) reallocation of workers across occupations. The existence of welfare gains through the first channel is unsurprising: workers have insufficient assets needed to span the risks they face. Still, this is the classic channel considered for analyzing welfare gains from social insurance and we formally show its existence within our model in the next proposition. To do so, we fix the labor allocation and output from the competitive equilibrium and provide an alternative feasible consumption allocation that Pareto dominates the competitive allocation.

Proposition 3.3 (The Competitive Allocation of Consumption Without Insurance is Pareto Inefficient). *Let $\{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ satisfy Definition 3.1 for the case $b =$*

---

[8]To explain the notation, it is easy to show directly that this corresponds to the complete markets equilibrium, in Appendix B.1

$\tau = 0$. *There exists an alternative feasible allocation* $\{\hat{c}_1(j), \hat{c}_d(j), \hat{c}_n(j), \hat{\ell}(j)\}$ *that:*
(i) increases expected utility in each occupations

$$E_j[U(\hat{c}_1(j), \hat{c}_n(j), \hat{c}_d(j))] \geq E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))] \; \forall j \in [0, J]$$
$$\exists k \; s.t. \; E_k[U(\hat{c}_1(k), \hat{c}_n(k), \hat{c}_d(k))] > E_k[U(c_1^*(k), c_n^*(k), c_d^*(k))]$$

(ii) is feasible

$$\int_j \hat{\ell}(j) \left(\hat{c}_1(j) + \theta(j)\hat{c}_d(j) + (1 - \theta(j))\hat{c}_n(j) \right) dj \leq \left(\int_j (\hat{\ell}(j)(2 - \theta(j)))^\gamma dj \right)^{\frac{1}{\gamma}}$$

Proof. See Appendix B.2 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

The next proposition formalizes the existence of the second welfare improving channel in our model: reallocation of labor to improve production efficiency and increase output. With constant-elasticity of substitution (CES) production (indeed, even with linear, $\gamma = 1$), efficient production requires the marginal product to be constant across occupations. This implies a constant life-time income across occupations. However, with incomplete markets, risk adverse workers require a wage premium to work in more risky occupations relative to less risky. For the competitive allocation to provide this risk premium, fewer workers must choose risky occupations than the efficient, output maximizing allocation. In other words, the competitive allocation of labor across occupations is first-order stochastically dominated by the efficient allocation.

Proposition 3.4. *[The competitive allocation without insurance puts too few workers in risky occupations] Let* $\ell^*(j)$ *satisfy Definition 3.1 for the case* $b = \tau = 0$. *Let* $\ell^{cm}(j)$ *be the feasible, output-maximizing allocation. Then,*

$$\int_{j=0}^t \ell^*(j)dj \leq \int_{j=0}^t \ell^{cm}(j)dj \quad \forall t \in (0, J]$$

This is to say, the efficient distribution of labor across occupations first-order stochastically dominates the distribution in the competitive allocation.

Proof. See Appendix B.4 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

The degree of productive inefficiency in the competitive equilibrium depends on both the extent of risk aversion of workers and the elasticity of substitution across occupations in production. As risk aversion increases, the competitive allocation becomes less efficient by concentrating even more workers in less-risky occupations. Comparably, if workers are risk neutral, the competitive allocation is efficient. Figure 2 illustrates this for constant relative risk aversion (CRRA) preferences given by: $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $\gamma = 0.5$.
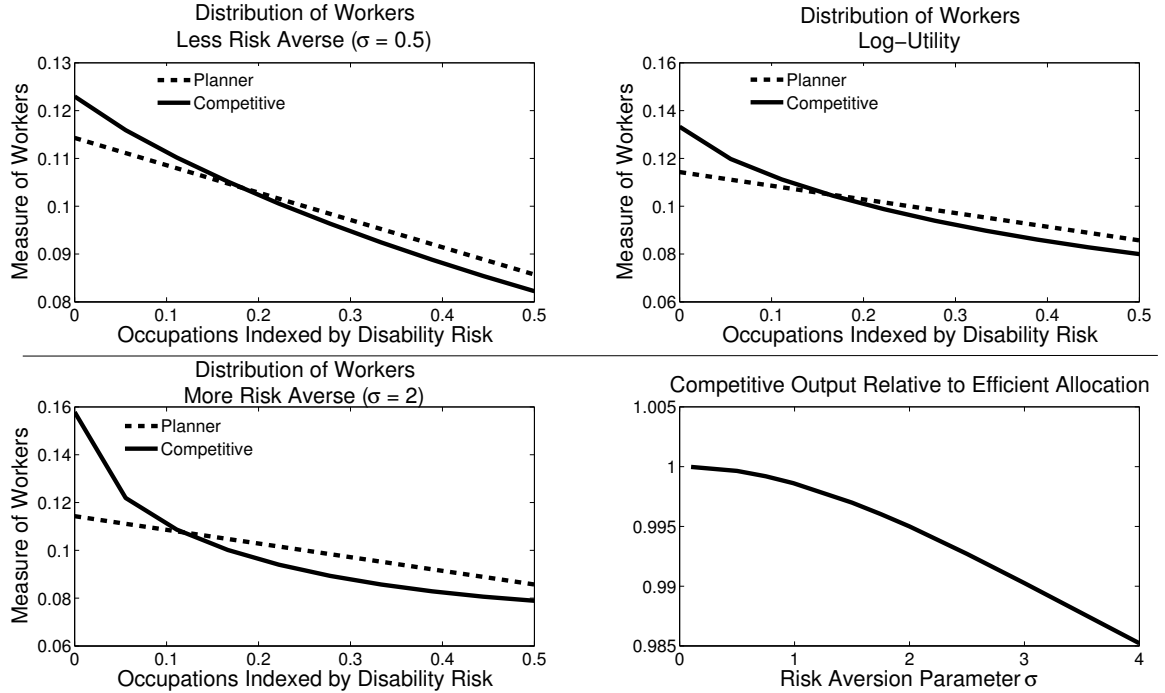

Figure 2: The competitive allocation relative to the efficient, planner's at vaious levels of risk aversion. The labor allocation diverges more greatly from the optimal planner's at higher levels of risk aversion.

As the elasticity of substitution between occupations increases, both the competitive and the efficient allocations place fewer workers in the risky occupations. While the efficient allocation always places more workers in the risky occupations, the distance between the two distributions is non-monotone. When occupations are perfect substitutes, the distributions are equal: Both allocations place all of the workers in the safest occupation. When occupations are perfect complements (production is Leontief), both allocations evenly distribute workers across occupations. Figure 3 illustrates this, fixing risk aversion at $\sigma = 2$.
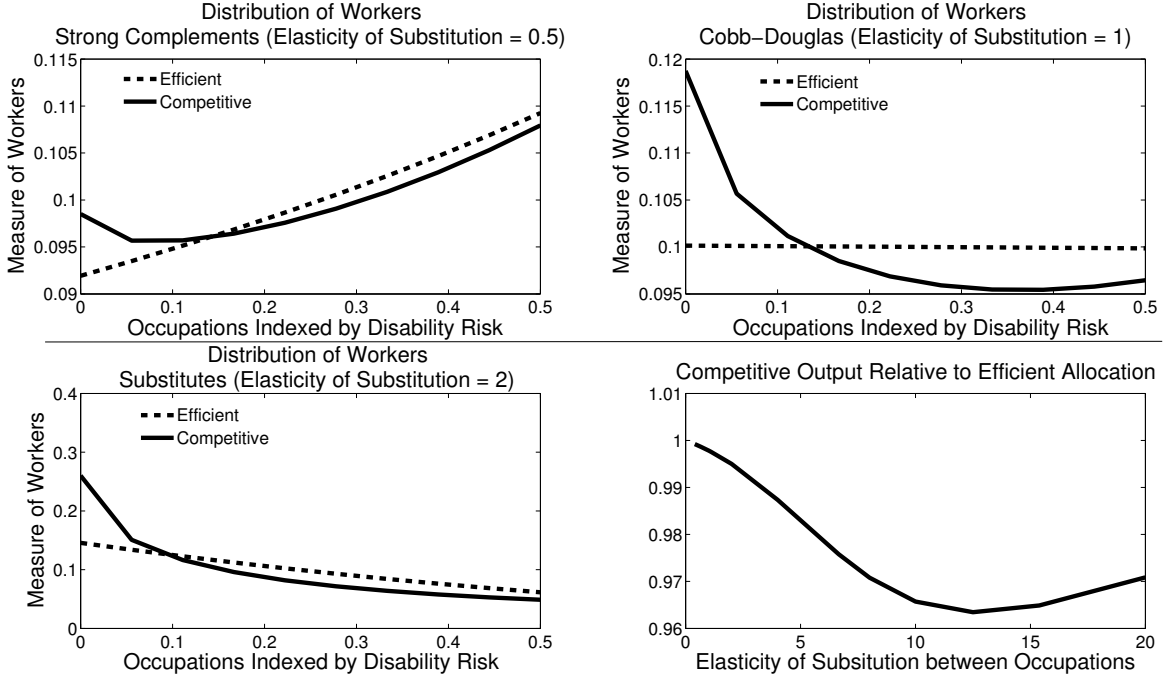
Figure 3: The competitive allocations relative to the efficient at various elasticities of substitution. Distortions away from the optimal planner's increase in γ to a point, then at infinity they again converge. The optimal allocation places more in the risky occupatiion when they are gross complements and more in the safe when they are substitues.

3.2. How Social Disability Insurance Improves Welfare

We have shown that the competitive equilibrium without an SDI scheme is inefficient, both in terms of inadequate consumption smoothing and misallocation of labor across occupations. SDI addresses both of these. It directly provides smoother consumption in risky occupations which indirectly serves to incentivize workers to choose riskier occupations. The next proposition shows that a marginal introduction of SDI improves both welfare and productive efficiency. It shifts labor to a more efficient allocation and provides smoother consumption across states.

Proposition 3.5 (From $\tau = 0$, insurance is Pareto improving on the margin)**.** *Let $EU(j, \tau)$ be the expected utility $E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))](\tau)$ from the allocation in competitive equilibrium $\{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ satisfying Definition 3.1 for occupation-independent*

benefit rate b funded by occupation-independent tax rate τ.

$$\left.\frac{\partial EU(j,\tau)}{\partial \tau}\right\|_{\tau=0} > 0 \quad \forall j$$

Proof. See Appendix B.5. □

A key insight is that the redistribution of workers across occupations under social insurance increases the wage in the risk-free occupation. This is a necessary condition for SDI to provide a Pareto improvement, regardless of specific parametrizations. Because workers are free to move across occupations, their welfare must equalize. In particular, there must be welfare gains in the risk-free occupation that does not benefit from a transfer in a disabled state. This is possible because each dollar of benefits offsets the wage premium required in risky occupations for self-insurance by more than one-to-one.

Welfare gains from expanding the SDI system are not without bounds. As the system becomes larger, it serves as an inefficient transfer to individuals in risky occupations. At the extreme of 100% taxation, the SDI system is clearly is clear suboptimal as every worker will choose the occupation that makes claiming disability insurance most likely. The following corollary shows that there exists and optimal SDI system between zero and this extreme.

Corollary 3.6 (There exists an welfare maximizing level of insurance). $\exists \tau^* > 0$ *such that*

$$\frac{\partial \int \ell(j) EU(j,\tau^*) dj}{\partial \tau} = 0$$

The welfare maximizing level of social insurance does not maximize output if occupation specific benefits do not exist. This is because a system that simultaneously maximizes output and welfare requires consumption to be perfectly smoothed across states in each occupation. In such a system there is no longer any risk in any occupation, workers then reallocate to equalize the marginal product across occupations, which is the efficient allocation. Without occupation specific benefits, the policy maker simply does not have sufficient policy tools to maintain indifference across occupations at the output maximizing allocation.

Proposition 3.7 (The welfare maximizing level of social insurance with occupation independent taxes and benefits does not maximize output.). *Let $n^{cm}(j)$ characterize the efficient (output maximizing) allocation. Let $n^{rp}(j), w^{rp}(j); \tau, b$ be the constrained optimal planner allocation (maximizes welfare in competitive equilibrium given limited policy tools). Then $y^{rp} < y^{cm}$ and, in particular $\int_{j=0}^{t} n^{rp}(j) dj < \int_{j=0}^{k} n^{cm}(j)$ for all $k \in [0, J]$.*

Proof. See Appendix B.7. □

As a Corollary to Proposition 3.7, we show the first best, planner's allocation can be achieved if occupation specific benefits are available.

Corollary 3.8 (The first-best planner allocation can be achieved with a lump-sum or proportional tax and occupation-specific benefits.)**.** *Let* $\mathcal{A}^*(\hat{b}, \hat{\tau}) = \{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ *satisfying Definition 3.1 given arbitrary occupation-specific benefits* $\hat{b} = \{b_j\}_j$ *funded by occupation-specific taxes* $\hat{\tau} = \{\tau_j\}_j$. *Let* $\mathcal{A}^{cm} = \{c_1^{cm}(j), c_n^{cm}(j), c_d^{cm}(j), a^{cm}(j), n^{cm}(j), \ell^{cm}(j)\}$ *define the first-best planner's allocation. Then,* $\exists \mathbf{b}, \tau$ *such that* $\mathcal{A}^*(\hat{b}, \hat{\tau}) = \mathcal{A}^{cm}$

Proof. See Appendix B.7 □

Figure 4 shows features of welfare and output gains from the occupation independent tax and transfer social insurance system described in the propositions above. First, starting with no social insurance, the welfare and output gains are always positive. Second, there exists a unique output-maximizing tax level and a unique welfare-maximizing tax level. The welfare maximizing tax level does not coincide with the output maximizing tax level. For a range of parametrizations ($\sigma = 3$ shown), the welfare-maximizing system puts more labor in risky occupations than the output-maximizing (efficient) allocation. This means that at the output-maximizing allocation, there are greater gains from providing higher benefits for consumption smoothing than the output loss incurred as these benefits move more workers into risky occupations than is efficient.

In our model with homogeneous workers, SDI increases ex ante welfare uniformly because of indifference across occupations. Ex post, we can see more welfare changes relative to no SDI for workers in each state: disabled and non-disabled. To understand these welfare effects, consider an old generation who entered the labor market without any SDI program and hold fixed their asset choices. We then give them the wages, taxes, and benefits of various sized SDI programs. The welfare gains for this experiment, across occupations and disability outcome, are shown in Figure 5.

The first panel in Figure 5 shows that disability benefits improve the welfare of workers that become disabled. The second panel shows workers with low disability risk benefit even if they do not become disabled. This is because they earn higher wages because of reallocation. Similarly, workers with high disability risk are actually worse off if they do not
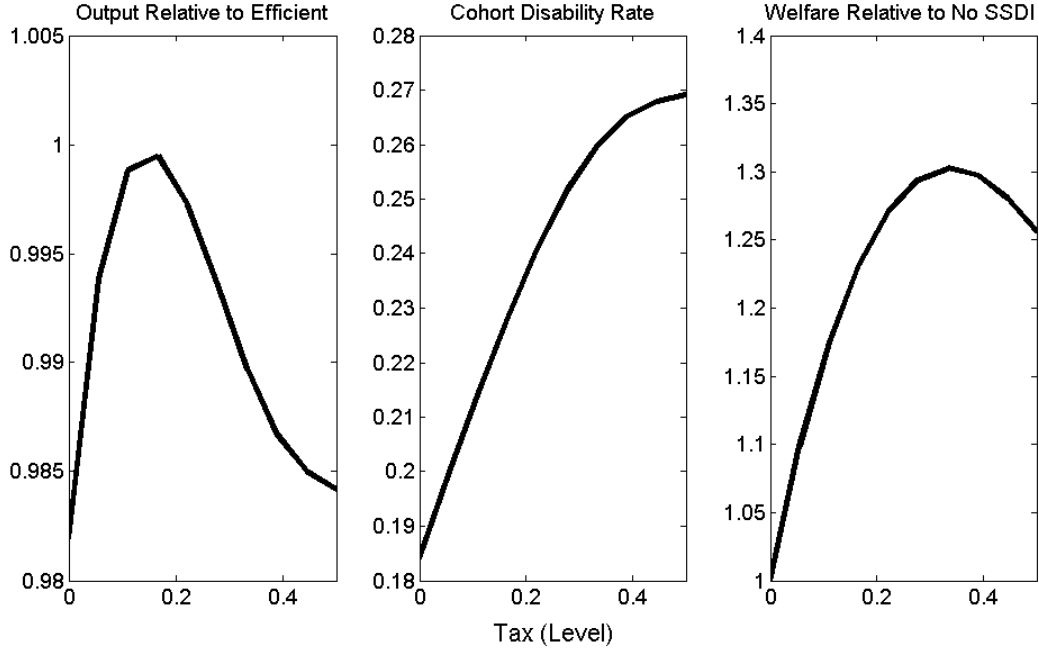
19

Figure 4: Output, disability, and welfare for different tax levels. The peak output is achieved at a lower tax than the welfare maximizing tax.

become disabled. This is because SDI causes more workers to enter the high risk occupations and lowers wages for workers in these occupations, but they gain in an ex ante sense due to risk sharing. The final panel shows the total expected benefit for workers in the second period of life.

3.3. Robustness and Extensions

In this section we show the main results of the baseline model, that the introduction of SDI increases welfare through each channel of risk sharing and occupational reallocation, are robust to considerations specific to the literature on disability risk as opposed to income risk generally. We separately consider three potential extensions: heterogeneity in patience, costly disability, and non-verifiable disability status. Then, we show how public, *social* disability insurance dominates private contracts within our framework.

Heterogeneity and Sorting. A main concern in our empirical work was to control for the sorting of individuals into occupations based on fixed, ex-ante heterogeneity. There, we were
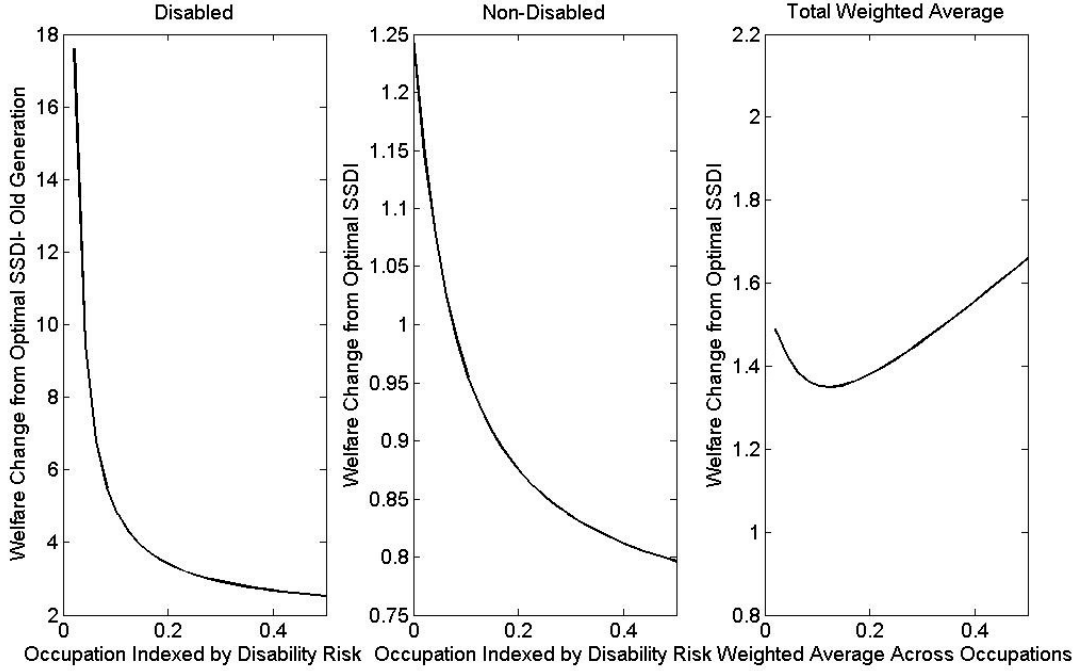
20

Figure 5: The welfare effect of SDI on different groups. Among low-risk occupations, the benefit comes from increased output and therefore wages, which is seen in ex post gains. High-risk occupations gain in an ex ante sense from risk sharing.

concerned such a fixed trait could affect occupation choice and health. Here we investigate within our model how such a fixed trait could affect how individuals sort across occupations and how they value disability insurance. For exposition purposes, we consider heterogeneity in discount factor, β and a two-type case, $\beta^H = 1$ and $0 < \beta^L = \beta < 1$. Our results are general to N types and adaptable to other forms of ex ante heterogeneity. Supporting proofs are in Appendix Section B.8.

Our first result shows monotone occupational sorting, a cut-off rule.[9] Patient, high-beta

---

[9]Sorting is also consistent with findings in our empirical work where the IV-implied effects were greater than OLS, which suggests sorting into high risk occupations. Our IV focused especially on tolerance to physical demands, but our conclusions from preference heterogeneity mostly extend to heterogeneity in health-sensitivity, though the model is less tractable.

types choose to occupy all of the less-risky occupations below an endogenous threshold $\bar{j}(\tau)$. Impatient low-beta types occupy the remaining riskier occupations. Our second result is that the marginal introduction of SDI increases welfare *for both types* through the same channels, risk sharing and occupational allocation. The first-best planner's labor allocation is the same as without preference heterogeneity and achieves higher output than the economy without SDI by putting more labor in risky occupations. A marginal introduction of SDI moves the labor allocation towards more risky occupations within beta-types and also moves the threshold $\bar{j}(\tau)$ of occupations held by high-beta types upwards. Both reallocations increase output.

An interesting result of this extension is that workers in risky occupations can have lower welfare gains from introducing SDI than workers in less risky occupations. As in Schulhofer-Wohl (2011), some agents value insurance less than others, here it is the low-beta types. These low-beta types also gain the least from the second channel, occupational reallocation, because as the threshold $\bar{j}(\tau)$ increases, the patient workers move into the relatively high-paying jobs. So, while the marginal introduction of SDI benefits the impatient types, further increases make them worse off.Individuals with low and even zero disability risk value SDI more than individuals with high risk and this is a qualitative distinction from models of SDI that treat disability risk as common across individuals.

Costly Disability. Disability is different from other types of income risk in that the decline in income associated with disability occurs in conjunction with an increase in expenditure costs. These costs range from direct medical costs to increased expenditure on goods and services that an individual can no longer produce themselves. We introduce such costs into the model here, with technical details in Appendix Section B.9. Let preferences over consumption in the disabled state be represented by the utility function u^d. We assume:

$$(i) \quad u_d(c(1+\chi)) = u(c) \quad ; \quad (ii) \quad u'_d(c(1+\chi)) \propto u'(c) \quad ; \quad \chi \geq 0, \qquad \forall c > 0.$$

That is the cost of disability is a constant proportion of consumption $(1 + \chi)$ required to regain the utility of the non-disabled. This is analogous to the preferences used in Low et al. (2015). Costs of disability have constant marginal relationships with the level of consumption and we can isolate concerns with productive efficiency and risk-sharing separately.

With this modification, the social planner's problem becomes:

$$\max_{\ell(j),c_d(j),c_n(j),c_1(j)} \int_j [u(c_1(j)) + \theta(j)u^d(c_d(j)) + (1-\theta(j))u(c_n(j))]\ell(j)dj$$

$$\text{s.t.} \int_j \ell(j)[c_1(j) + \theta(j)c_d(j) + (1-\theta(j))c_n(j)]dj \leq \left(\int_j (2-\theta(j))^\gamma \ell(j)^\gamma dj\right)^{\frac{1}{\gamma}}, \quad \int_j \ell(j)dj = 1$$

The solution to this problem delivers two main results. First, the allocation of consumption equalizes marginal utilities across disabled and non-disabled individuals: $c = c_1 = c_n = (1+\chi)c_d$. Second, the labor in occupation k relative to occupation j at the optimal allocation of labor is:

$$\frac{\ell(k)}{\ell(j)} = \left(\frac{(2-\theta(k))}{(2-\theta(j))}\right)^{\frac{\gamma}{1-\gamma}} \left(\frac{(2+\chi\theta(k))}{(2+\chi\theta(j))}\right)^{\frac{1}{1-\gamma}}$$

With $\chi = 0$, this is equivalent to the base case without a cost of disability. We see the disability cost, $\chi > 0$ reduces labor in the high-risk occupation. The relative marginal-product is now $\frac{2-\theta(j)}{2+\chi\theta(j)}/\frac{2-\theta(k)}{2+\chi\theta(k)}$: $\chi > 0$ drives a wedge between the relative marginal product of labor between two occupations of different risk. Increasing this cost χ increases the difference in the marginal product of labor between two occupations for all values of γ, even strong complements.

In a sense, this is a second cost of disability in our model. The first cost was the "human capital" cost of fewer workers given by the fraction $\theta(j)$ who cannot work in the second period. The key difference is that the impact of the human capital cost on the planner's allocation depends on the elasticity of substitution across occupations. For $\gamma > 0$, the occupations are gross complements and the planner actually puts more labor in risky occupations relative to safe ones. The qualitative impact of the utility costs captured by χ does not depend on this elasticity; these costs always decrease relative labor in risky occupations. However, the magnitude of the effect is increasing in the substitutability of occupations.

We provide a set of propositions that show the introduction of SDI improves welfare through both the risk sharing channel and the channel of labor reallocation that increases aggregate productivity. These results, however, do not hold for arbitrary costs χ. A sufficient restriction is that the disability cost must not exceed the elasticity of the benefit replacement rate b to the tax τ. Such an assumption ensures that a cost of τc in the non-disabled state is offset by a benefit of $\frac{b}{1+\chi}c$ in the disabled state.

Non-verifiable Disability Risk and Status. We consider social disability insurance if neither ex-ante disability risk or realized disability status is unobservable. As one would expect, to induce truth-telling, benefits cannot be too high or at least one type of worker would choose to falsely report herself as disabled. With our setup, this is guaranteed if the replacement rate is less than $\frac{1}{2}$. Proposition B.16 makes this result explicit.

In our baseline model, we have already assumed that disability risk, $\theta(j)$ is non-verifiable, and therefore SDI does not discriminate based on occupation. Now, we show that under the optimal SDI scheme, no worker will misreport her status and provide an analytical bound such that truth-telling is always incentive compatible.

If the replacement rate b is less than $1 - \tau$, given that benefits and taxes are both proportional to wages a non-disabled agent clearly gains from working, and therefore will not misreport her status.[10] With $b < \frac{1}{2}$ we can guarantee this holds and numerically verify it in the calibration, Section 4.

Private Contracts. In the US, private disability insurance is limited to a small payout and offered only to a subset of workers. Golosov and Tsyvinski (2006) do not consider private contracts because of this. We will show numerically how adverse selection hinders private disability insurance, providing a rationale for social insurance. Private insurance contracts induce too much demand from workers in high-risk occupations, sullying the risk pool. Private insurance contracts therefore earn negative profits, because at any price and benefit that would be feasible with compulsory SDI, low-risk agents opt not to buy. The market breaks down in a way typical of adverse selection.

The crucial constraint for private insurance is that, like social insurance, they cannot condition on the worker's occupational risk level, $\theta(j)$.[11] But, whereas social insurance can dictate the amount of insurance any agent takes, private insurance cannot. With social insurance, workers can only increase their exposure to the insurance by increasing their exposure to disability risk, which we showed was actually optimal to an extent (Proposition 3.5). Instead, with private insurance there is over-subscription into these contracts from the

---

[10]With a constant as the benefit there is another bound that is a function of the entire worker distribution.

[11]If either private or social insurance can verify disability risk level (given by occupation) then they can achieve the first best allocation. This is the case with complete markets (Section B.1, Equation B.16). For comparison's sake, we keep the same information restrictions on both social and private insurance.

workers already in the risky occupations. On the other hand, if occupational risk-level were verifiable and private insurance contracts on it, these can dominate social insurance. We do not explore that here further because relaxing a constraint for the private insurers makes the comparison uneven.

To illustrate how private disability insurance breaks down, we consider our baseline economy, fixing labor at its efficient allocation and then introducing private disability contracts. Given an occupation j and insurance price p, agents solve for optimal disability insurance g:

$$\max_{g,a} u(w^{CM}(j) - pw^{CM}(j)g - a) + \theta(j)u(bw^{CM}(j)g + a) + (1 - \theta(j))\, u(w^{CM}(j) + a)$$

Then the benefits b are set to satisfy the insurers' budget constraint if they could enforce uniform participation, the best-case scenario.

$$\int_j \theta(j)pw^{CM}(j) - (2 - \theta(j))bw^{CM}(j)dj = 0$$

At any price, the private insurance contracts incur negative profits, as shown in the left pane of Figure 6 because those in high-risk occupations over subscribe, as shown in the right pane. The insurer could increase p or reduce b, but both of these serve only to worsen the selection problem. Figure 6 shows how increases in p lead to increases in the average $\theta(j)$ among the insured population and the effect is analogous if the insurer dropped b. For private contracts with positive p and b, profits are negative. Social disability insurance can mandate broad participation across risk levels, but private disability insurance fails here because of adverse selection: the workers who choose to participate most are the worst risks.

4. Quantitative Evaluation

We now calibrate our theoretical model to suggest the relative magnitudes of reallocation and risk-sharing. Within the policy region of the US-calibrated economy and the constrained optimal, labor allocation changes very little, instead increased risk sharing is the dominant force.

For this exercise, we maintain the two-generation overlapping generations structure of the theory section but modify the duration of the first period to 30 years (ages 25-55) and the second period to 10 years (ages 55-65). With an annual time discount rate of 0.95, the second period is discounted to $\beta = 0.103$. We consider only two occupations, guided by the
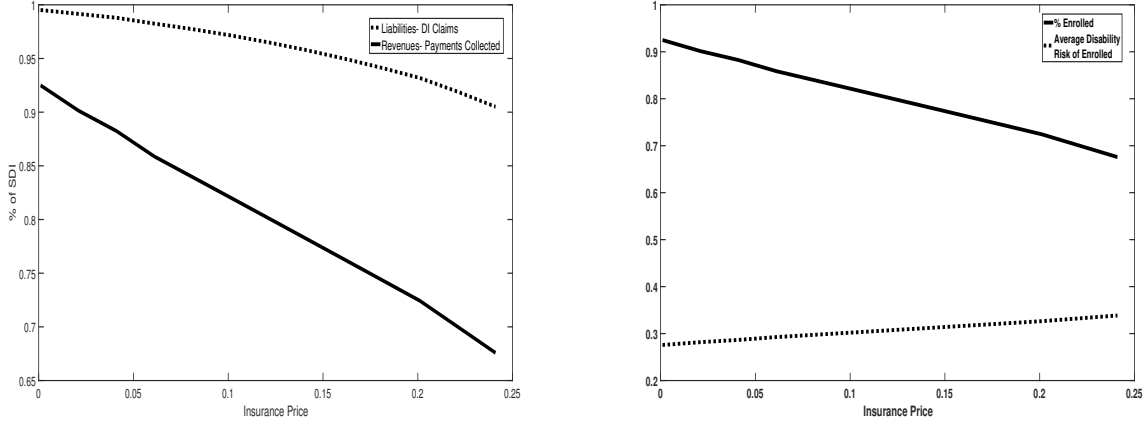
Figure 6: As the price of disability insurance increases, the contracts remain unprofitable (Left) and the risk pool deteriorates (Right). On left, Revenue and outlays ofcontracts are normalized to their fracion of SDI with the same price (tax).

natural break in occupational risks seen in Figure 1. The base risk in the low-risk occupation is an 8% probability of disability before retirement and the additional risk in the high-risk occupations amounts to a 16% probability of disability before retirement.

We give the social disability system a proportional tax and replacement rate and the same tools to our constrained optimal planner. The baseline replacement rate to 40% (see Duggan and Autor (2006)) and then solve for the implied balanced-budget tax.

Preferences are $u(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$ with $\sigma = 2$. The production technology uses low-disability-risk labor n_ℓ and high-risk labor n_h: $Y = Q(\alpha n_\ell^\gamma + (1-\alpha)n_h^\gamma)^{1/\gamma}$. We compute results for an elasticity of substitution ranging from $\frac{1}{3}$ to 10, recalibrating each time. Specifically, we calibrate α to match 20% of workers choosing the high-risk occupation and then set Q to normalize output to 1.

Using the calibrated production, preference, and occupational risk parameters, we then calculate two different insurance policies: the constrained optimal equilibrium and an equilibrium with "actuarially fair" insurance, defined by benefits that cover the expected income loss for the average worker.

Table 3 presents aggregate statistics for each of these three policy regimes. Relative to the actuarilly fair system, the constrained optimal, constrained optimal SDI system is a substantial improvement in welfare, 2.6% in consumption equivalent terms. But the constrained

26

optimal policy only increases welfare over the baseline by 0.2% despite rather large changes in policy, the planner's replacement rate is 47.5% larger than the baseline calibration.

	Elasticity of Substitution		
	$\frac{1}{3}$	1	2.5
Young Labor in Risky Occupation			
Baseline	20%	20%	20%
Constrained Optimal	20.05%	20.09%	20.23%
Actuarially Fair	19.78%	19.56%	18.91%
Welfare Gain from Constrained Optimal DI			
(consumption equiv rel. to Baseline)	+0.2%	+0.2%	+0.2%
(consumption equiv rel. to Actuarially Fair)	+2.6%	+2.6%	+2.6%
Constrained Optimal Output			
(rel. to Baseline)	0.99	0.99	0.99
(rel. to Actuarially Fair)	1.00	1.00	0.99
Aggregate Disability			
(rel. to baseline)	+0.05%	+0.08%	+0.19%
(rel. to Actuarially Fair)	+1.35%	+2.76%	+14.5%
Replacement rate			
Baseline	40.00%	40.00%	40.00%
Constrained Optimal	58.90%	58.88%	58.81%
Tax rate			
Baseline	0.99%	0.99%	0.99%
Constrained Optimal	1.45%	1.45%	1.45%

Table 3: The optimal system compared to observed US system, constrained optimal policy, and actuarially fair in a calibrated model

Notice however, that labor changes very little from the baseline to the constrained optimal policies. Essentially, all of the welfare gains there come from additional risk sharing. To further investigate this, Figure 7 shows the labor allocation under each policy while changing the elasticity of substitution. The baseline is calibrated to put 20% in the risky occupation, which is more than the output-maximizing allocation for each γ. The constrained optimal SDI system puts even more workers in the risky occupation than the baseline calibration.

But, for a reasonable elasticity of substitution, the allocation in the risky occupation is always within a few percent of the baseline calibration. We conclude that the baseline calibration already captures the output gains from the reallocation of labor. Further welfare gains of the constrained optimal allocation come primarily from increased risk-sharing.
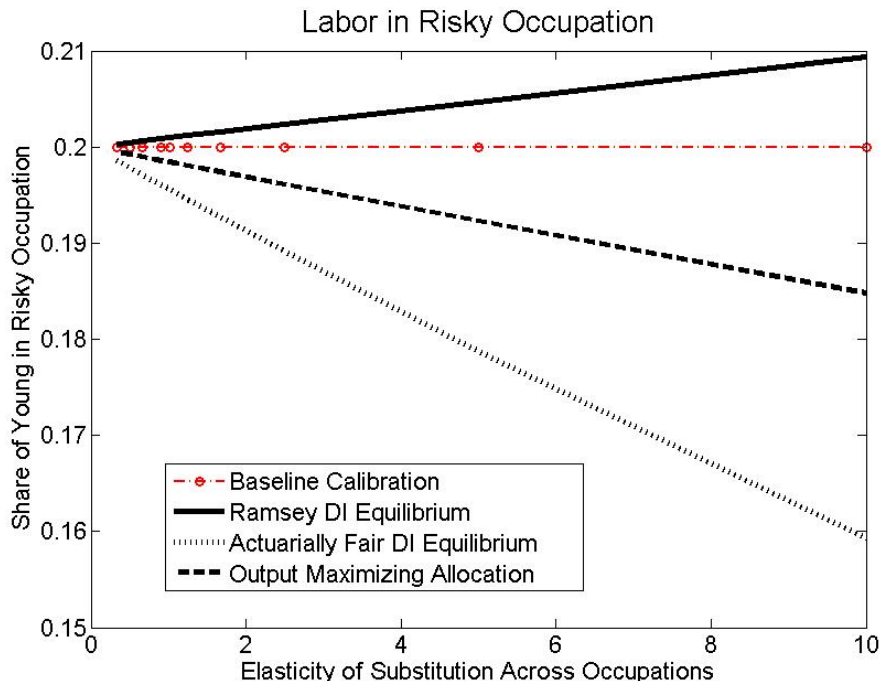
Figure 7: Labor in the risky occupation is higher under the baseline and constrained optimal SDI programs than the output maximizing allocation. It is lower with only actuarially fair DI.

5. Conclusion

In this paper, we have explored how occupational choice shapes disability risk and policy. We first documented the importance of lifetime occupational exposure to differences in disability risk. We then embedded this idea—occupational choices imply different levels of disability risk—into an equilibrium model with incomplete asset markets and imperfectly substitutable occupations. Here, incomplete markets for disability risk lead to both imperfect risk-sharing *and* an inefficient allocation of labor across occupations. This leaves room at the margin for a welfare-improving disability insurance, which improves consumption smoothing

and reallocates workers to increase output. This latter point is essentially moral hazard: insuring risky occupations encourages more risk-taking, but at the margin this is actually efficient.

Finally, we took quantitative insights from a two-occupation model calibrated to resemble the United States. We found that the welfare-maximizing SDI provides welfare gains that are equal to a 2.6% increase in consumption in a world with actuarially fair insurance alone. Between policy regimes, there was relatively little difference in labor allocation for most

For future research, our results on private insurance indicated potentially new avenues. Data on private disability insurance is not inconsistent, but is more complicated than our model admits: Private disability insurance is very rarely sold directly to individuals, in 2015 there were about a half-million in the US (Isenberg (2016)). Employer-provided disability insurance is instead much more common, covering about $\frac{1}{3}$ of the workforce. This coverage, however, is essentially rationed and occupation is a good indicator of access. 98% of workers enroll if DI is offered by their employers, but offers are very heterogeneous across occupations (Groshen and Perez (2015)). [12] Management and professional occupations, relatively low risk, are offered long-term disability insurance 50% of the time, while high risk work, production occupations and construction or extraction occupations are only offered it 28% and 20% of the time, respectively. Employer provision of disability insurance seems to be one way private insurance can condition on $\theta(j)$, by utilizing employers' additional information and then rationing insurance. The National Compensation Survey has very detailed information on DI access within and across firms and understanding its patterns could be enriched by incorporating worker choice across occupation and firm.

From a quantiative side, our analysis has abstracted from potentially important details of the SSDI application process and how they interact with the rich dynamics of health. Our model was intentionally parsimonious to focus our point: Disability risk differs across occupations, and when social disability insurance affects worker's choice of risk, it also affects occupational choice. Once workers have suffered some health consequences, when they decide to apply or not is potentially important.

---

[12]This large heterogeneity across occupations in access to disability insurance is, a useful validation of our empirical work, showing evidence that private markets treat occupations differently in what private disability insurance exists.

List of references

Acemoglu, D., Shimer, R., October 1999. Efficient unemployment insurance. Journal of Political Economy 107 (5), 893–928.

Administration, S. S., 2013. Annual statistical report on the social security disability insurance program, 2012. Social Security Administration, Office of Retirement and Disability Policy: Office of Research, Evaluation and Statistics.

Duggan, M., Autor, D., 2006. The growth in the social security disability rolls: a fiscal crisis unfolding. Tech. Rep. 12436, National Bureau of Economic Research.

Fletcher, J. M., Sindelar, J. L., Yamaguchi, S., June 2009. Cumulative effects of job characteristics on health. Working Paper 15121, National Bureau of Economic Research.

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

Groshen, E. L., Perez, T. E., September 2015. National compensation survey: Employee benefits in the united states, march 2015. Bulletin 2782, U.S. Bureau of Labor Statistics.

Gruber, J., 1997. The consumption smoothing benefits of unemployment insurance. The American Economic Review, 192–205.

Health and Retirement Study, 2013. HRS Core public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI.

Isenberg, K. N., 2016. U.s. individual disability income insurance 2015 annual supplement. Tech. rep., LIMRA.

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

Low, H., Meghir, C., Pistaferri, L., 2015. Disability insurance and the dynamics of the incentive-insurance tradeoff. American Economic Review.

Morefield, G. B., Ribar, D. C., Ruhm, C. J., February 2011. Occupational status and health transitions. Working Paper 16794, National Bureau of Economic Research.

RAND HRS Data, Version P, August 2016. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA.

Ravesteijn, B., van Kippersluis, H., van Doorslaer, E., Sep. 2013. The Wear and Tear on Health: What is the Role of Occupation? MPRA Paper 50321, University Library of Munich, Germany.

Rust, J., Phelan, C., July 1997. How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets. Econometrica 65 (4), 781–832.

Schulhofer-Wohl, S., 2011. Heterogeneity and Tests of Risk Sharing. Journal of Political Economy 119 (5), 925 – 958.

Online, not-for-publication appendix

A. Empirical appendix

In this Appendix, we include details about our instrumental variables approach using O*NET measures and robustness using an alternative disability definition.

A.1. Instrumental Variables

Table 4 presents the first stage regression results for the endogenous variable H_j, health requirements in occupation j. Each non-physical requirement is standardized to have a mean 0 and standard deviation of 1. The requirements are defined by the first 3 orthogonal principal components of O*NET descriptors excluding physical requirements.

O*NET Non-phys Reqs, PC1	-0.712 **	(0.00211)
O*NET Non-phys Reqs, PC2	0.269 **	(0.00335)
O*NET Non-phys Reqs, PC3	0.302 **	(0.00261)
Constant	-0.00901 **	(0.00249)
Observations	14763	
R^2	0.933	
F	107126.6	

Robust standard errors in parentheses

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 4: First stage estimates using non-physical descriptors to instrument physical requirements.

A.2. Robustness: Industry-clustering and fixed effects

In this section we include industry-level fixed effects and cluster standard errors on these industries. We define a worker's industry analogously to our definition of occupation, using the industry associated with the worker's longest-held job. There are 13 industry categories,

1

similar to the level of aggregation of occupations.[13] The industry-level fixed effects change the point estimates little. Clustering standard errors approximately doubles them, but our key coefficients are all still strongly signficant and instruments still cannot be rejected as valid. Table 5 presents our main regression results.

Table 6 presents the first stage regression results for the endogenous variable H_j, health requirements in occupation j. Each non-physical requirement is standardized to have a mean 0 and standard deviation of 1. The requirements are defined by the first 3 orthogonal principal components of O*NET descriptors excluding physical requirements.

A.3. Robustness: self-reported health limitations to work

To check the robustness that occupations differ in their affect on disability, instead of using the presence of an ADL difficulty, we might use a self-reported limitation to work based on health grounds. This figure tells almost exactly the same story. Figure 8 shows this distribution, constructed in the same manner as the previous two distributions. Again, the tail of high-risk occupations is quite long and the large high-risk occupations are transportation and machine operators with more than 200% above the clustering of lower-risk occupations.
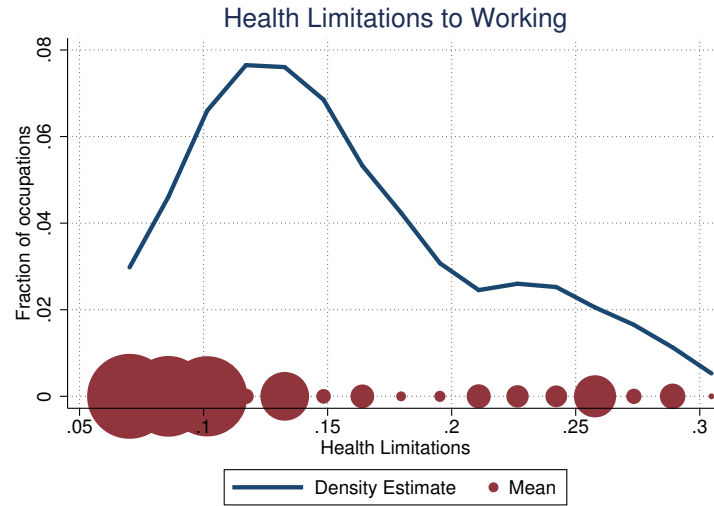

Figure 8: The density across occupations of the rate reporting a health limitation to work

---

[13] A list of names is available in the public documentation for Health and Retirement Study (2013).

As in our analysis with ADL difficulty, we can use a Oaxaca-Blinder decomposition to separate occupation-effects and effects from observables in the difference in risk between the high and low risk occupations. Consistent with the previous decomposition on the probability of an ADL limitation, we use the same regressors: a cubic for potential experience, body mass index (BMI), time, dummies for education level, gender, marital status, race, and tobacco use. Again, the occupation effect is large, though differences in observables are a slightly more important factor in the difference in self-reported health limitations to work.

Finally, we include our IV regression estimates using health limitations as the report of disability. The regression results in Table 8. The marginal effects are about 40% as strong as with ADL difficulty to measure disability but all of the relationships are still very statistically significant.

B. Theoretical Results and Proofs

B.1. Setting up the tax-free economy

We first define the complete and incomplete markets versions of our tax-free economy ($\tau = b = 0$). In both cases, firms operate in spot markets (no long-term contracting). The representative firm solves

$$\max_{\{n(j)\}} \left(\int_{j=0}^{J} n(j)^\gamma \right)^{1/\gamma} dj - \int_{j=0}^{J} w(j)n(j)dj$$

In equilibrium all non-disabled workers are hired: $n(j) = (2 - \theta(j))\ell(j)$. From this problem, the first-order condition (FOC) is

$$w(j) = \left(\frac{y}{n(j)} \right)^{1-\gamma} \tag{B.1}$$

And for any occupations j, k

$$\frac{w(j)}{w(k)} = \left(\frac{n(j)}{n(k)} \right)^{\gamma-1} \tag{B.2}$$

Complete Markets. In this section, we identify the endogenous variables of the competitive equilibrium with the superscript *cm*. We begin by setting up the households' problem with the complete set of Arrow securities along with a backyard storage technology

earning return R_0. It is trivial to show directly that these allocations correspond to those in the social planner's problem because the absence of discounting combined with a backyard storage technology solves typical incomplete markets inefficiencies associated with OLG economies. With the full set of assets, both welfare theorems hold and the efficient allocation, $\{c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j), \ell^{cm}(j)\}$, is also unique.Their possible states are $s = \{1, d, n\}$ for period 1, 2 as disabled, and 2 as non-disabled; and j for the occupation they choose. The full set of Arrow securities, $\{a_d(j), a_n(j)\}$ spans states j, s:

$$\max_{\{c_s(j), \ell(j), a_n(j), a_d(j)\}} \int_{j=0}^{J} \{u\left(c_1(j)\right) + \theta(j)u\left(c_d(j)\right) + (1 - \theta(j))u\left(c_n(j)\right)\} \ell(j)dj \tag{B.3}$$

$$c_1(j) \leq w(j) - a(j) - \left(\int_{k=0}^{J} a_n(k) + a_d(k)\right) dk \tag{B.4}$$

$$c_n(j) \leq w(j) + R_0 a(j) + \int_k R_n(k)a_n(k)dk \tag{B.5}$$

$$c_d(j) \leq R_0 a(j) \int_k R_d(k)a_d(k)dk \tag{B.6}$$

$$1 \geq \int_j \ell(j)dj \tag{B.7}$$

The solution to this problem sets each interest rate $R_n(j) = \frac{1}{1-\theta(j)}$, $R_d(j) = \frac{1}{\theta(j)}$ and the saving policies are $a_n^{cm}(j) = \frac{-\theta(j)}{2}(1 - \theta(j))w^{cm}(j)$; and $a_d^{cm}(j) = \frac{2-\theta(j)}{2}\theta(j)w^{cm}(j)$ for j such that $\ell(j) = 1$ and $a_s^{cm}(k) = 0 \ \forall k \neq j$. This implies consumption is smoothed across time and states

$$c_1^{cm}(j) = c_n^{cm}(j) = c_d^{cm}(j) = \frac{2 - \theta(j)}{2}w^{cm}(j)$$

The occupation choice $\ell(j)$ requires expected earnings to be equalized across occupations j, k:

$$w^{cm}(j)(2 - \theta(j)) = w^{cm}(k)(2 - \theta(k)) \tag{B.8}$$

Combining the wage condition with consumption gives

$$w^{cm}(j) = \frac{2c}{2 - \theta(j)} \tag{B.9}$$

where c is the consumption level of every individual. From the firms' side, Equation B.1, and substituting market clearing that $y = 2c$, and $\ell^{cm}(j)(2 - \theta(j)) = n(j)$, we also have

$$w^{cm}(j) = \left(\frac{2c}{(2 - \theta(j))\ell^{cm}(j)}\right)^{1-\gamma} = \frac{2c}{2 - \theta(j)}$$

4

Solving for $\ell^{cm}(j)$ gives

$$\ell^{cm}(j) = \left(\frac{2c}{2 - \theta(j)}\right)^{\frac{\gamma}{\gamma - 1}} \tag{B.10}$$

Now, we can use the market clearing condition that $\int_j \ell^{cm}(j)dj = 1$ to solve for c, and hence $\ell^{cm}(j)$. That is, $1 = \left(\int_j \frac{2c}{2-\theta(j)}dj\right)^{\frac{\gamma}{\gamma-1}}$, which, solving for c, implies that

$$c = \frac{1}{2}\left(\int_{j=0}^J (2 - \theta(j))^{\frac{\gamma}{1-\gamma}}dj\right)^{\frac{1-\gamma}{\gamma}}$$

and then plugging that in:

$$\ell(j) = (2 - \theta(j))^{\frac{\gamma}{1-\gamma}}\left[\int_{k=0}^J (2 - \theta(k))^{\frac{\gamma}{1-\gamma}}dk\right]^{-1} \tag{B.11}$$

Social Planner. Notice that the complete markets allocation will yield the same labor allocation as the social planner and hence the same output. Because consumption is also equal across all states *and generations*, the social planner and complete markets allocations are the same and the complete markets allocation is efficient.

The social planner solves

$$\max_{\ell(j),c(j)} \int_j \lambda(j)2u(c(j))\ell(j) \text{ s.t.} \tag{B.12}$$

$$\left(\int_j ((2 - \theta(j))\ell(j))^\gamma \, dj\right)^{1/\gamma} \geq \int_j l(j)2c(j)dj \tag{B.13}$$

$$\int_j \ell(j)dj \leq 1 \tag{B.14}$$

where we have already plugged in the obvious result that the social planner provides households with perfectly smooth consumption, setting it to be equal in both periods and all states. $\lambda(j)$ defines arbitrary Pareto weights.

The FOC on labor implies that

$$\left(\frac{y^{cm}}{\ell^{cm}(j)(2 - \theta(j))}\right)^{1-\gamma}(2 - \theta(j)) = \left(\frac{y^{cm}}{\ell^{cm}(k)(2 - \theta(k))}\right)^{1-\gamma}(2 - \theta(k)) \tag{B.15}$$

for any j, k where $y^{cm} = \left(\int_j ((2 - \theta(j))\ell^{cm}(j))^\gamma\right)^{1/\gamma}$. Then, express all $\ell^{cm}(k)$ in terms of $\ell^{cm}(j)$ and we can use the condition on labor, Equation B.14 to sum up:

$$1 = \int_k \left(\frac{\ell(j)^{cm\ \gamma-1}(2 - \theta(j))^\gamma}{(2 - \theta(k))^\gamma}dk\right)^{\frac{1}{\gamma-1}}$$

5

Solving gives

$$\ell^{cm}(j) = (2 - \theta(j))^{\frac{\gamma}{1-\gamma}} \left[\int_k (2 - \theta(k))^{\frac{\gamma}{1-\gamma}} dk \right]^{-1} \tag{B.16}$$

Note that this is the same outcome as our complete markets setup and hence output is the same in both worlds. Because consumption is equal for all people in all states in both frameworks, we have shown that the allocations are equivalent.

Incomplete Markets. With only one asset, households solve

$$\max_{j,a} u(w(j) - a) + \theta u(a) + (1 - \theta(j))u(w(j) + a) \tag{B.17}$$

$$\{a\} : u'(w^*(j) - a^*(j)) = \theta u'(a^*(j)) + (1 - \theta(j))u'(w^*(j) + a^*(j)) \tag{B.18}$$

$$\{j\} : u(w^*(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j) + a^*(j))$$
$$= u(w^*(k) - a^*(k)) + \theta(k)u(a^*(k)) + (1 - \theta(k))u(w^*(k) + a^*(k)) \tag{B.19}$$

The FOC on occupation choice that households are indifferent between choosing any risk level $\theta(j), \theta(k)$. We have denoted the optimal savings policy of an agent choosing j as $a^*(j)$, which is distinct from the complete markets case where we indexed potential $a_s(k)$ by the level of risk assumed and the occupation that would buy them.

Consider the case $\theta(k) = 0$, then the two first-order-conditions imply

$$u(w^*(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j) + a^*(j)) = 2u(w^*(k)) \tag{B.20}$$

Note also that the Inada condition on utility guarantees that for $j > 0$, $\theta(j) > 0$ implies $a^*(j) > 0$ otherwise in the disabled state marginal utility is not finite.

B.2. Inefficiency of the incomplete markets equilibrium
Proof of 3.3 . (reprinted here): Let $\{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ satisfy Definition 3.1 for the case $b = \tau = 0$. There exists an alternative feasible allocation $\{\hat{c}_1(j), \hat{c}_d(j), \hat{c}_n(j), \hat{\ell}(j)\}$ such that

$$E_j[U(\hat{c}_1(j), \hat{c}_n(j), \hat{c}_d(j))] \geq E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))] \; \forall j \in [0, J]$$
$$\exists k \; s.t. \; E_k[U(\hat{c}_1(k), \hat{c}_n(k), \hat{c}_d(k))] > E_k[U(c_1^*(k), c_n^*(k), c_d^*(k))]$$

where E_s conditions on disability probability θ_s; and

$$\int_j \hat{\ell}(j) \left(\hat{c}_1(j) + \theta(j)\hat{c}_d(j) + (1 - \theta(j))\hat{c}_n(j) \right) dj \leq \left(\int_j (\hat{\ell}(j)(2 - \theta(j)))^\gamma \right)^{\frac{1}{\gamma}}$$

6

Proof. It is easy to construct one such, Pareto dominating allocation to the incomplete markets competitive allocation. Define $\hat{\ell}(j) = \ell^*(j)$ and $\hat{w}(j) = w^*(j)$ for all j, the same values as the competitive equilibrium. Next set $\hat{R}_d(j) = \frac{1}{\theta(j)}, \hat{R}_n(j) = \frac{1}{1-\theta(j)}$ $\forall j$. Define the remaining allocations as the solution to agents' problem taking prices $\{\hat{w}(j)\}$ and $\{\hat{R}_n(j), \hat{R}_d(j)\}$ as given:

$$\max_{c_1(j),c_n(j),c_d(j),a_d(j),a_n(j)} u(\hat{w}(j) - a_d(j) - a_n(j)) + \theta(j)u(a_d(j)\hat{R}_d(j)) + (1-\theta(j))u(\hat{w}(j) + a_n(j)\hat{R}_n(j))$$

We get fully smooth consumption consumption, $\hat{c}_1(j) = \hat{c}_n(j) = \hat{c}_d(j) = \frac{2-\theta(j)}{2}\hat{w}(j)$. This is clearly feasible, as we have only redistributed the same output among workers who choose occupation j. This is strictly preferred for each $j \in (0, J]$ because of strict concavity and for $j = 0$, the workers in the risk-free occupation are indifferent. $\qquad\square$

B.3. Comparing the complete and incomplete markets allocations

Lemma B.1. *Compared to the incomplete markets allocation, the complete markets allocation has more mass in every risky occupation relative to the zero-risk occupation: $\forall j \in (0, J]$, $\frac{\ell(j)}{\ell(0)} < \frac{\ell^{cm}(j)}{\ell^{cm}(0)}$.*

Proof. To prove this, we use contradiction. Hence, we begin with the counter factual,[14]

$$\exists\, j \text{ s.t. } \frac{\ell(j)}{\ell(0)} \geq \frac{\ell^{cm}(j)}{\ell^{cm}(0)} \tag{B.21}$$

We can then take this to the wage space using Equation B.2. First multiply both sides, $\frac{(2-\theta(j))\ell(j)}{2\ell(0)} \geq \frac{(2-\theta(j))\ell^{cm}(j)}{2\ell^{cm}(0)}$. Raising both the power $\gamma - 1$ we get, subject to a parameter restriction that $\gamma < 1$

$$\left(\frac{(2-\theta(j))\ell(j)}{2\ell(0)}\right)^{\gamma-1} \leq \left(\frac{(2-\theta(j))\ell^{cm}(j)}{2\ell^{cm}(0)}\right)^{\gamma-1}$$

This implies, by Equation B.2

$$\frac{w(j)}{w(0)} \leq \frac{w^{cm}(j)}{w^{cm}(0)} \tag{B.22}$$

---

[14]It is trivial to show the case of $\theta(0) = 0$ because $\frac{\ell(0)}{\ell(0)} = 1 = 1 = \frac{\ell^{cm}(0)}{\ell^{cm}(0)}$. Hence, we prove for cases in which $\theta(j) \in (0, J]$

Next, from the occupation choice indifference condition, workers must be indifferent between either occupation

$$u(w(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))(w(j) + a^*(j)) = 2u(w(0))$$

And, by Equation B.22 and monotonicity

$$2u\left(w(0)\right) \geq 2u\left(\frac{w^{cm}(0)}{w^{cm}(j)}w(j)\right)$$

Then, because $\frac{w^{cm}(0)}{w^{cm}(j)} = \frac{2 - \theta(j)}{2}$ as shown in Equation B.8, we get that

$$2u\left(\frac{w^{cm}(0)}{w^{cm}(j)}w(j)\right) = 2u\left(w(j)\frac{2 - \theta(j)}{2}\right)$$

where the right-hand side is the expected earnings from occupation θ. Combining these,

$$u(w(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))(w(j) + a^*(j)) = 2u(w(0)) \geq 2u\left(w(j)\frac{2 - \theta(j)}{2}\right)$$

But this means that the incomplete markets, risky allocation is weakly preferred to the expected earnings, which violates Jensen's inequality with strictly concave preferences. Hence, our assumption in Equation B.21 could not be and we have established our Lemma B.1. $\square$

Corollary B.2. *There is more mass at the risk-free occupation, where $\theta(0) = 0$, in the incomplete markets allocation with $\tau = 0$ than the efficient allocation: $\ell^*(0) > \ell^{cm}(0)$*

Proof. This is easy to see, using the fact that both ℓ^*, ℓ^{cm} are densities and hence $\int_j \ell^*(j) = \int_j \ell^{cm}(j) = 1$.

Taking inequality $\frac{\ell^*(j)}{\ell^*(0)} < \frac{\ell^{cm}(j)}{\ell^{cm}(0)}$ $\forall j > 0$ and integrating over j we have

$$\int_j \frac{\ell^*(j)}{\ell^*(0)}dj < \int_j \frac{\ell^{cm}(j)}{\ell^{cm}(0)}dj \Leftrightarrow \frac{\int \ell^*(j)dj}{\int \ell^{cm}(j)dj} < \frac{\ell^*(0)}{\ell^{cm}(0)}$$

Then because $\int_j \ell^*(j)dj = \int_j \ell^{cm}(j) = 1$ by definition, this gives us $1 < \frac{\ell^*(0)}{\ell^{cm}(0)}$

$\square$

B.4. Inefficiency of competitive incomplete markets allocation

Proof of 3.3. (reprinted here): *Let* $\{c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j), \ell^{cm}(j)\}$ *be the efficient, planner's allocation solving*

$$\max_{\{c_1(j),c_d(j),c_n(j),\ell(j)\}} \int_j \ell(j) \left(u(c_1(j)) + \theta(j)u(c_d(j)) + (1 - \theta(j))u(c_n(j)) \right) dj \text{ s.t.}$$

(B.23)

$$\left(\int_j ((2 - \theta(j))\ell(j))^\gamma dj \right)^{\frac{1}{\gamma}} \geq \int_j \ell(j) \left(c_1(j) + \theta(j)c_d(j) + (1 - \theta(j))c_n(j) \right) dj$$

(B.24)

$$1 \geq \int_j \ell(j)dj$$

(B.25)

Then $\{c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j), \ell^{cm}(j)\}$ *strictly Pareto dominates the competitive equilibrium with zero taxes* $\{c_1^*(j), c_d^*(j), c_n^*(j), \ell^*(j)\}$

Proof. First, note that both allocations satisfy feasibility because of the problems from which they are defined. That is

$$\int_j \ell(j) \left(c_1(j) + \theta(j)c_d(j) + (1 - \theta(j))c_n(j) \right) dj \leq \left(\int_j (\ell(j)(2 - \theta(j)))^\gamma dj \right)^{\frac{1}{\gamma}}$$

for both $\{c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j), \ell^{cm}(j)\}$, $\{c_1^*(j), c_d^*(j), c_n^*(j), \ell^*(j)\}$

Then, from Corollary B.2 $\ell^*(0) > \ell^{cm}(0) \Leftrightarrow w^*(0) < w^{cm}(0)$ and because the utility function is strictly increasing $2u(w^*(0)) < 2u(w^{cm}(0))$. Adding in the occupational choice indifference conditions gives, for arbitrary j:

$$u(w^*(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j) + a^*(j)) = 2u(w^*(0))$$
$$< 2u(w^{cm}(0))$$
$$= 2u(w^{cm}(j)\frac{2 - \theta(j)}{2})$$

or equivalently, using the indifference conditions:

$$E_j[U(c_1^*(j), c_d^*(j), c_n^*(j))] = E_0[U(c_1^*(0), c_d^*(0), c_n^*(0))]$$
$$< E_j[U(c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j))]$$
$$= E_0[U(c_1^{cm}(0), c_d^{cm}(0), c_n^{cm}(0))]$$

9

This means that $\forall j$, $E_j[U(c_1^{cm}(j), c_d^{cm}(j), c_n^{cm}(j))] > E_j[U(c_1^*(j), c_d^*(j), c_n^*(j))]$ just as we required.

$\square$

Lemma B.3. *The ratio of labor in any two occupations, one riskier than the other, is greater in the complete markets allocation than the equilibrium allocation with incomplete markets in which $b = \tau = 0$:*

$\forall j, k$ *such that* $\theta(j) < \theta(k)$, $\frac{\ell^{cm}(k)}{\ell^{cm}(j)} > \frac{\ell^*(k)}{\ell^*(j)}$; *these distributions satisfy the monotone likelihood ratio property.*

Proof. First, suppose this is not the case, then we assume

$$\frac{\ell^{cm}(k)}{\ell^{cm}(j)} \leq \frac{\ell^*(k)}{\ell^*(j)} \tag{B.26}$$

Which is equivalent to

$$\frac{\ell^{cm}(k)}{\ell^*(k)} \leq \frac{\ell^{cm}(j)}{\ell^*(j)}$$

and by the same reasoning as in Lemma B.1 (along with parameter restriction $\gamma < 1$)

$$\frac{w^*(k)}{w^{cm}(k)} \leq \frac{w^*(j)}{w^{cm}(j)}$$

And because we have established that $\frac{w^{cm}(k)}{w^{cm}(j)} = \frac{2-\theta(j)}{2-\theta(k)}$ that gives us that that

$$w^*(k)(2 - \theta(k)) < w^*(j)(2 - \theta(j)) \tag{B.27}$$

This is to say, the expected income in occupation j is higher than that in k and $\theta(k)$ also has more risk. The lottery associated with $\theta(j), w(j)$ dominates that of lottery $\theta(k), w(k)$. Clearly, this cannot be consistent with household's indifference.

To put this more formally, consider the sequence of inequalities:

$$u(w^*(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j) + a^*(j)) \tag{B.28}$$

$$> u(w^*(j) - a^*(k)) + \theta(j)u(a^*(k)) + (1 - \theta(j))u(w^*(j) + a^*(k)) \tag{B.29}$$

$$\geq u(w^*(k) - a^*(k)) + \theta(k)u(a^*(k)) + (1 - \theta(k))u(w^*(k) + a^*(k)) \tag{B.30}$$

This is a contradiction of the household-indifference condition across occupations. The first holds by definition of $a^*(j), a^*(k)$ via the Maximum Theorem and the second holds because of

our contradictory assumption. However, by our occupational-choice indifference condition, B.20,

$$u(w^*(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j) + a^*(j))$$
$$= u(w^*(k) - a^*(k)) + \theta(k)u(a^*(k)) + (1 - \theta(k))u(w^*(k) + a^*(k))$$

Therefore, $\ell*$ and ℓ^{cm} have the monotone likelihood ratio property, with ℓ^{cm} dominating ℓ^*

$\square$

3.4]Proof of 3.4 . (reprinted here): Let $\ell^*(j)$ satisfy Definition 3.1 for the case $b = \tau = 0$. Let $\ell^{cm}(j)$ be the feasible, output-maximizing allocation. Then,

$$\int_{j=0}^t \ell^*(j)dj \leq \int_{j=0}^t \ell^{cm}(j)dj \quad \forall t \in (0, J]$$

The efficient distribution of labor across occupations first-order stochastically dominates the distribution in the competitive allocation.

Proof. This is an application of the monotone likelihood ratio property shown in Lemma B.3. Given that the distributions have the monotone likelihood ratio property and $\ell^*(j)$ is dominated by $\ell^{cm}(j)$, this implies that $\int_{j=0}^t \ell^*(j)dj \leq \int_{j=0}^t \ell^{cm}(j)dj \quad \forall t \in (0, J]$

$\square$

B.5. Proof that the introduction of SDI is Pareto Improving ($\tau > 0$ is welfare improving)

Proof. Proof of 3.5 (reprinted here): [From $\tau = 0$, insurance is Pareto improving on the margin] Let $EU(j, \tau)$ be the expected utility $E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))](\tau)$ from the allocation in competitive equilibrium $\{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ satisfying Definition 3.1 for occupation-independent benefit rate b funded by occupation-independent tax rate τ.

$$\frac{\partial EU(j, \tau)}{\partial \tau}\bigg|_{\tau=0} > 0 \quad \forall j$$

We will show that $\exists j, \delta$ such that $\forall \epsilon \in (0, \delta) \; EU(j, \epsilon) - EU(j, 0) > 0$: It must be the case that $\exists j$ such that total earnings are higher than with zero taxes. Then, if at least one agent

has higher earnings, that agent is better off and because of indifference across occupations, so too is everyone else.

We have already shown in Lemma B.7 that $\exists \delta$ such that $\forall \epsilon \in (0, \delta)$, output increases, $y(\epsilon) > y(0)$. For any ϵ within this ball, we can re-write this in terms of income:

$$\int_{j=0}^{J} \ell^*(j, \epsilon) w(j, \epsilon) \left((2 - \theta(j))(1 - \epsilon) + \theta(j) b(\epsilon) \right) dj > \int_{j=0}^{J} \ell^*(j, 0) w(j, 0) \left((2 - \theta(j)) \right) dj$$

This inequality might exist for one of two reasons: (1) $w(j, \epsilon) \left((2 - \theta(j)) + \theta(j) b(\epsilon) \right) > w(j, 0) \left((2 - \theta(j)) \right)$ for at least some j, or (2) $\ell^*(j, \epsilon)$ and $\ell^*(j, 0)$ are different but $\nexists j$ such that expected income is higher.

Suppose (2) holds, that is:

$$w(j, \epsilon) \left((2 - \theta(j))(1 - \epsilon) + \theta(j) b(\epsilon) \right) \le w(j, 0) \left(2 - \theta(j) \right) \ \forall j \tag{B.31}$$

If this is the case then we can multiply both sides of this inequality by $\ell^*(j, \epsilon)$

$$\ell^*(j, \epsilon) w(j, \epsilon) \left((2 - \theta(j))(1 - \epsilon) + \theta(j) b(\epsilon) \right) \le \ell^*(j, \epsilon) w(j, 0) \left(2 - \theta(j) \right) = w(j, 0) n^*(j, \epsilon) \tag{B.32}$$

Because this holds for every j, we will integrate over j and the inequality becomes

$$\int \ell^*(j, \epsilon) w(j, \epsilon) \left((2 - \theta(j))(1 - \epsilon) + \theta(j) b(\epsilon) \right) dj \le \ell^*(j, \epsilon) w(j, 0) \left(2 - \theta(j) \right) = \int w(j, 0) n^*(j, \epsilon) dj$$

The left-hand side is exactly output with $\tau = \epsilon$ and we can replace $w(j, 0)$ with the solution to the firms' problem $w(j, 0) = y(0)^{1-\gamma} n^*(j, 0)^{\gamma-1}$.

$$y(\epsilon) \le y(0)^{1-\gamma} \int_{j} n^*(j, \epsilon) n^*(j, 0)^{\gamma-1} dj$$

We can manipulate the right-hand side:

$$y(\epsilon) \leq y(0)^{1-\gamma} \int_j n^*(j,\epsilon) n^*(j,0)^{\gamma-1} dj$$

$$= y(0)^{1-\gamma} \int_j \frac{n^*(j,\epsilon)}{n^*(j,0)} n^*(j,0)^{\gamma} dj$$

$$= y(0)^{1-\gamma} \int_j \frac{n^*(j,\epsilon)/n^*(0,0)}{n^*(j,0)/n^*(0,0)} n^*(j,0)^{\gamma} dj$$

$$\leq y(0)^{1-\gamma} \int_j \frac{n^*(j,\epsilon)/n^*(j,0)}{n^*(j,0)/n^*(0,0)} n^*(j,0)^{\gamma} dj$$

$$\leq y(0)^{1-\gamma} \int_j n^*(j,0)^{\gamma} dj$$

$$= y(0)^{1-\gamma} y(0)^{\gamma} = y(0)$$

Where the second and third to last come from $\frac{n^*(j,\epsilon)/n^*(0,0)}{n^*(j,0)/n^*(0,0)} \leq \frac{n^*(j,\epsilon)/n^*(0,\epsilon)}{n^*(j,0)/n^*(0,0)} \leq 1$, which we showed in Lemma B.5.

The result $y(\epsilon) \leq y(0)$, contradicts Lemma B.7. Thus, it must be the case that $\exists j$ such that $w(j,\epsilon)\left((2-\theta(j)) + \theta(j)b(\epsilon)\right) > w(j,0)\left((2-\theta(j))\right)$

In case (1), because we have imperfect credit markets (only a storage technology, no borrowing or private insurance), it is not immediate that higher expected lifetime consumption makes workers better off. Consider the problem of a worker in occupation j with some tax rate $\epsilon \in (0,\delta)$ for which $w(j,\epsilon)[(2-\theta(j))(1-\epsilon) + \theta(j)b(\epsilon)] > w(j,0)(2-\theta(j))$. If $j = 0$ ($\theta_0 = 0$) it is obvious that $EU_0(\epsilon) > EU_0(0)$. For all other j, we have

$$w(j, 1-\epsilon) > w(j,0) - \frac{\theta(j)}{2-\theta(j)} b(\epsilon) w(j,\epsilon)$$

This implies for ϵ sufficiently small:

$$EU_j(\epsilon) > max_a u(w(j,0) - \frac{\theta(j)}{2-\theta(j)} b(\epsilon) w(j,\epsilon) - a)$$

$$+ \theta(j) u(b(\epsilon)w(j,\epsilon) + a) + (1-\theta(j))u(w(j,0) - \frac{\theta(j)}{2-\theta(j)} b(\epsilon)w(j,\epsilon) + a)$$

Then, for the marginal introduction of $b(\tau)w(j,\tau)$ we have can differentiate with respect to the benefit $b(\epsilon)w(j,\epsilon)$ at $\epsilon = 0$

$$\frac{dEU_j(0)}{d(bw(j,0))} > -\frac{\theta(j)}{2-\theta(j)}u'(w(j,0)-a^*(j,0)) + \theta(j)u'(a^*(j,0)) - \frac{\theta(j)}{2-\theta(j)}(1-\theta(j))u'(w(j,0)+a^*(j,0))$$

$$= \theta(j)u'(a^*(j,0)) - \frac{\theta(j)}{2-\theta(j)}[u'(w(j,0)-a^*(j,0)) + (1-\theta(j))u(w(j,0)+a^*(j,0))]$$

$$> \theta(j)[u'(a^*(j,0)) - u'(w(j,0)-a^*(j,0))]$$

$$> 0$$

where $a^*(j,0)$ is the optimizing policy at $bw(j,\tau) = 0$. The second inequality comes from $u(\cdot)$ being concave. The final inequality comes from $a^*(j,0) < w(j,0) - a^*(j,0)$ since $\theta(j) \in [0,\frac{1}{2}]$, which implies $u'(a^*(j,0)) > u'(w(j,0) - a^*(j,0))$. □

B.6. Supporting lemmas for $\tau > 0$ is welfare improving

Lemma B.4. *For wages $w(j)$, benefits $b(\tau)$ and tax rate τ and some occupational choices $\ell(j)$ and wages $w(j)$, the benefit is greater than the tax rate: $b(\tau) > \tau$.*

Proof. Given the tax rate τ, the government budget constraint holds that

$$b(\tau)\int_{j=0}^{J}\theta(j)w(j)\ell(j)dj = \tau\int_{j=0}^{J}(2-\theta(j))w(j)\ell(j)dj$$

Then for $\theta(j) < 2 - \theta(j)$, the whole valid domain of θ, $b > \tau$ □

Lemma B.5. *Given the competitive equilibrium allocation $\{c_1^*(j,\tau), c_n^*(j,\tau), c_d^*(j,\tau), a^*(j,\tau), n^*(j,\tau), \ell^*(j,\tau)\}$ with tax rate τ,*

$$\frac{\partial}{\partial\tau}\left(\frac{\ell^*(j,\tau)}{\ell^*(0,\tau)}\right)\bigg|_{\tau=0} > 0$$

Proof. To sign the derivative, we will show that $\exists\delta$ such that $\forall \epsilon \in (0,\delta)$, $\left(\frac{\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)}\right) > \left(\frac{\ell^*(j,0)}{\ell^*(0,0)}\right)$. This δ can be anything small enough such that $a^*(j,\delta) > 0$ for every j. That is to say, the borrowing constraint does not bind and choices are all in the interior of the budget set.

Holding wage constant, for $j > 0$ the expected income at $\tau = 0$ and $\epsilon > 0$ is $(2-\theta(j))w(j,0)$ and $(2-\theta(j))w(j,0)(1-\epsilon) + \theta(j)\hat{b}(\epsilon|0)w(j,0)$. We define $b(\epsilon)$ as the level of benefits implied by $\tau = \epsilon$ and $w(j,\epsilon)$.

The expected income $(2 - \theta(j))w(j,0) < (2 - \theta(j))w(j,0)(1 - \epsilon) + \theta(j)\hat{b}(\epsilon)w(j,0)$ because of Lemma B.4 and therefore, by the Maximum Theorem, expected utility is higher:

$$\max_{a_\epsilon} u(w(j,0)(1 - \epsilon) - a_\epsilon) + \theta(j)u(w(j,0)b(\epsilon) + a_\epsilon) + (1 - \theta(j))u(w(j,0)(1 - \epsilon) + a_\epsilon)$$
$$> \max_{a_0} u(w(j,0) - a_0) + \theta(j)u(a_0) + (1 - \theta(j))u(w(j,0) + a_0)$$

For $j = 0$, however, the expected utility is lower with the ϵ tax rate:

$$2u(w(0,0)(1 - \epsilon)) < 2u(w(0,0))$$

Based on indifference under $\tau = 0$, we have

$$\max_{a_\epsilon} u(w(j,0)(1 - \epsilon) - a_\epsilon) + \theta(j)u(w(j,0)b(\epsilon) + a_\epsilon) + (1 - \theta(j))u(w(j,0)(1 - \epsilon) + a_\epsilon)$$
$$> 2u(w(0,0)(1 - \epsilon)) \tag{B.33}$$

But indifference with $\tau = \epsilon$ implies that $EU(j,\epsilon) = EU(0,\epsilon)$. Therefore, $w(j,\epsilon) \neq w(j,0)$ and $\frac{w(j,\epsilon)}{w(0,\epsilon)} \neq \frac{w(j,0)}{w(0,0)}$

We now suppose the contradiction to our premise, that $\exists \hat{j}$ such that

$$\frac{\ell^*(\hat{j},\epsilon)}{\ell^*(0,\epsilon)} \leq \frac{\ell^*(\hat{j},0)}{\ell^*(0,0)} \Leftrightarrow \frac{w(\hat{j},\epsilon)}{w(0,\epsilon)} \geq \frac{w(\hat{j},0)}{w(0,0)}$$

As we have just shown in Equation B.33, it cannot be that $\frac{w(\hat{j},\epsilon)}{w(0,\epsilon)} = \frac{w(\hat{j},0)}{w(0,0)}$, therefore

$$\frac{w(\hat{j},\epsilon)}{w(0,\epsilon)} > \frac{w(\hat{j},0)}{w(0,0)}$$

But then, $\exists k$ such that $k > j$ but

$$\frac{w(\hat{j},\epsilon)}{w(0,\epsilon)} > \frac{w(k,\epsilon)}{w(0,\epsilon)} > \frac{w(\hat{j},0)}{w(0,0)}$$

which means that $w(\hat{j},\epsilon) > w(k,\epsilon)$ and therefore

$$(2 - \theta(\hat{j}))w(\hat{j},\epsilon)(1 - \epsilon) + \theta(\hat{j})b(\epsilon)w(\hat{j},\epsilon) > (2 - \theta(k))w(k,\epsilon)(1 - \epsilon) + \theta(k)b(\epsilon)w(k,\epsilon)$$

But if k is a riskier occupation with lower expected income than $\hat{j}$, then workers cannot be indifferent. For any such $\hat{j}$ we could find a k and therefore $\forall j$:

$$\frac{w(j,\epsilon)}{w(0,\epsilon)} < \frac{w(j,0)}{w(0,0)} \Leftrightarrow \frac{\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)} < \frac{\ell^*(j,0)}{\ell^*(0,0)}$$

$\square$

15

Corollary B.6. *The mass in the risk-free occupation, where $\theta(0) = 0$, is decreasing in τ:* $\left.\frac{\partial \ell^*(0,\tau)}{\partial \tau}\right|_{\tau=0} < 0$

Proof. This is a direct result of Lemma B.5 and the identity $\int_j \ell^*(j, \tau)dj = 1$. If $\exists \delta \; : \forall \epsilon \in (0, \delta) \; \frac{\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)} > \frac{\ell^*(j,0)}{\ell^*(0,0)}$. Consider integrating over j, then

$$\int_j \frac{\ell^*(j, \epsilon)}{\ell^*(0, \epsilon)} dj > \int_j \frac{\ell^*(j, 0)}{\ell^*(0, 0)} dj$$

$$\frac{\int_j \ell^*(j, \epsilon)dj}{\int_j \ell^*(j, 0)dj} > \frac{\ell^*(0, \epsilon)}{\ell^*(0, 0)}$$

$$1 > \frac{\ell^*(0, \epsilon)}{\ell^*(0, 0)}$$

$\square$

Lemma B.7. *Let, $y^*(\tau)$ be the output at competitive equilibrium allocation $\{c_1^*(j, \tau), c_n^*(j, \tau), c_d^*(j, \tau), a^*(j, \tau)n^*(j, \tau), \ell^*(j, \tau)\}$ with tax rate τ. $y^*(\tau)$ is increasing in τ at $\tau = 0$, $\frac{\partial y}{\partial \tau}|_{\tau=0} > 0$*

Proof. Recall the tax-free competitive equilibrium has lower output than the first best planner's efficient output, $y^*(0) \leq y^{cm}$ (as shown in 3.4) and that this is because the distribution of labor in the first best first order stochastically dominate the distribution of labor in the tax-free competitive equilibrium: $\int_{j=0}^{x} n^*(j, 0)dj \geq \int_{j=0}^{x} n^{cm}(j, 0)dj$ for all $x \in (0, J]$. Here, we use the related Lemma that helped establish monotone likelihood ratio dominance, B.3 and the similar Lemma relating $\tau > 0$ to $\tau = 0$, Lemma B.5. Using these, for any $j > 0$, we can find an $\upsilon(j) > 0$ small enough such that:

$$\frac{\ell^*(j, 0)}{\ell^*(0, 0)} < \frac{\ell^*(j, \upsilon(j))}{\ell^*(0, \upsilon(j))} < \frac{\ell^{cm}(j)}{\ell^{cm}(0)}$$

Let $\delta = \min_j \upsilon(j)$ so that $\forall \epsilon \in (0, \delta)$ and $\forall j$,

$$\frac{\ell^*(j, 0)}{\ell^*(0, 0)} < \frac{\ell^*(j, \epsilon)}{\ell^*(0, \epsilon)} < \frac{\ell^{cm}(j)}{\ell^{cm}(0)} \tag{B.34}$$

We now treat cases of $\gamma > 0$ and $\gamma < 0$ separately, though the steps are exactly the same. First, with $\gamma > 0$, we multiply Inequality B.34 by $(2 - \theta(j))$ and exponentiate to γ.

$$\left(\frac{(2 - \theta(j))\ell^*(j, 0)}{\ell^*(0, 0)} \right)^\gamma < \left(\frac{(2 - \theta(j))\ell^*(j, \epsilon)}{\ell^*(0, \epsilon)} \right)^\gamma < \left(\frac{(2 - \theta(j))\ell^{cm}(j)}{\ell^{cm}(0)} \right)^\gamma$$

16

Because this holds for all j, we integrate over j and then exponentiate to $\frac{1}{\gamma}$

$$\left(\int_j \left(\frac{(2-\theta(j))\ell^*(j,0)}{\ell^*(0,0)}\right)^\gamma dj\right)^{\frac{1}{\gamma}} < \left(\int_j \left(\frac{(2-\theta(j))\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)}\right)^\gamma dj\right)^{\frac{1}{\gamma}} < \left(\int_j \left(\frac{(2-\theta(j))\ell^{cm}(j)}{\ell^{cm}(0)}\right)^\gamma dj\right)^{\frac{1}{\gamma}}$$

$$\frac{y^*(0)}{\ell^*(0,0)} < \frac{y^*(\epsilon)}{\ell^*(0,\epsilon)} < \frac{y^{cm}}{\ell^{cm}(0)}$$

If $\gamma < 0$ we have the same, but inequalities flip twice when we exponentiate:

$$\frac{\ell^*(j,0)}{\ell^*(0,0)} < \frac{\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)} < \frac{\ell^{cm}(j)}{\ell^{cm}(0)}$$

$$\left(\frac{(2-\theta(j))\ell^*(j,0)}{\ell^*(0,0)}\right)^\gamma > \left(\frac{(2-\theta(j))\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)}\right)^\gamma > \left(\frac{(2-\theta(j))\ell^{cm}(j)}{\ell^{cm}(0)}\right)^\gamma$$

$$\left(\int_j \left(\frac{(2-\theta(j))\ell^*(j,0)}{\ell^*(0,0)}\right)^\gamma dj\right)^{\frac{1}{\gamma}} < \left(\int_j \left(\frac{(2-\theta(j))\ell^*(j,\epsilon)}{\ell^*(0,\epsilon)}\right)^\gamma dj\right)^{\frac{1}{\gamma}} < \left(\int_j \left(\frac{(2-\theta(j))\ell^{cm}(j)}{\ell^{cm}(0)}\right)^\gamma dj\right)^{\frac{1}{\gamma}}$$

$$\frac{y^*(0)}{\ell^*(0,0)} < \frac{y^*(\epsilon)}{\ell^*(0,\epsilon)} < \frac{y^{cm}}{\ell^{cm}(0)}$$

Now we have to use Lemma B.1 and Corollary B.6 to order $\ell^*(0,0), \ell^*(0,\epsilon)$ and $\ell^{cm}(0)$. This tells us $\ell^*(0,0) > \ell^*(0,\epsilon) > \ell^{cm}(0)$ and therefore

$$\frac{1}{\ell^*(0,0)} < \frac{1}{\ell^*(0,\epsilon)} < \frac{1}{\ell^{cm}(0)}$$

Therefore, we can re-write $\frac{y^*(0)}{\ell^*(0,0)} < \frac{y^*(\epsilon)}{\ell^*(0,\epsilon)}$ as $\ell^*(0,0)y^*(0) < y^*(\epsilon)\frac{\ell^*(0,0)}{\ell^*(0,\epsilon)}$ and because $y^*(\epsilon)\frac{\ell^*(0,0)}{\underset{<}{}}1$ this implies

$$y^*(0) < y*(\epsilon) .$$

□

B.7. Achieving the first best allocation

Proof. Proof of 3.7 (reprinted here): [The welfare maximizing level of social insurance with occupation independent tax and benefit rates does not maximize output.] Let $n^{cm}(j)$ characterize the efficient (output maximizing) allocation. Let $\{n^{rp}(j), w^{rp}(j)\}; \tau, b$ be the constrained optimal planner allocation (maximizes welfare in competitive equilibrium given policy tools). Then $y^{rp} < y^{cm}$.

We will show the competitive equilibrium indifference condition $E_j[U(w(j))] = E_0[U(w(0))]$ is not satisfied at the output maximizing allocation. Recall, wages at the output maximizing allocation satisfy:

$$w(j) = \frac{2}{2 - \theta(j)} w(0) \qquad \forall j$$

Therefore, we can write $E_j[U(w(j))]$ for arbitrary $j \in (0, J]$ as:

$$E_j[U(w(j))] = \max_a \{ u((1 - \tau - a) \frac{2}{2 - \theta(j)} w_0) +$$

$$+ \theta(j) u((b + a) \frac{2}{2 - \theta(j)} w_0) + (1 - \theta(j)) u((1 - \tau + a) \frac{2}{2 - \theta(j)} w_0)$$

By concavity of $u(\cdot)$, it must be $c_1 \leq E[c_2]$. Some algebra delivers:

$$c_1 \leq E[c_2] \Rightarrow (1 - \tau - a) \leq a + \theta(j) b + (1 - \theta(j))(1 - \tau)$$

$$\Rightarrow a \geq \frac{\theta(j)}{2} (1 - \tau - b)$$

$$\Rightarrow c_1 \leq w(0) \frac{2}{2 - \theta(j)} \frac{(2 - \theta(j))(1 - \tau) + \theta(j) b}{2}$$

$$\Rightarrow c_1 \leq w(0) [(1 - \tau) - \frac{\theta_j}{2 - \theta_j} b]$$

This inequality implies first period consumption in occupation j is less than that of the risk free occupation: $c_1^*(j) \leq (1 - \tau) w(0)$; if:

$$[(1 - \tau) - \frac{\theta_j}{2 - \theta_j} b] \leq (1 - \tau)$$

This is clearly the case since $\theta \in (0, J]$ and $b \geq 0$. Thus, we have shown $c_1(j) < c_1(0)$. Jensen's inequality and the Euler, $u'(c_1) = E[u'(c_2)]$ imply that $u(c_1) > E_j[u(c_2)]$ and together we have:

$$E_0[u(w(0))] = 2u(w(0)(1 - \tau)) \geq 2u(c_1(j)) > E_j[u(w(j))]$$

This contradicts the occupation indifference condition and so productive efficiency cannot be maintained as a competitive equilibrium when using a proportional tax and benefit scheme with occupation independent rates. $\qquad \square$

Proof. Proof of 3.8 (reprinted here): [The first-best planner allocation can be achieved with a lump-sum or proportional tax and occupation-specific benefits.] Let $\mathcal{A}^*(\widehat{b}, \widehat{\tau}) = \{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ satisfying Definition 3.1 given arbitrary occupation-specific benefits $b(j)$ funded by occupation-specific taxes $\tau(j)$. Let $\mathcal{A}^{cm} = \{c_1^{cm}(j), c_n^{cm}(j), c_d^{cm}(j), a^{cm}(j), n^{cm}(j)\ell^{cm}(j)\}$ define the first-best planner's allocation. Then, $\exists b(j), \tau(j)$ such that $\mathcal{A}^*(b((j)), \tau(j)) = \mathcal{A}^{cm}$

To achieve the efficient allocation in the decentralized economy, $w(j) = w(j')$ for all j, j' and $EU_j = EU_{j'}$. This can be achieved with occupation specific benefits and a non-occupation specific tax. These can be any combination of lump-sum or proportional. Here consider lump-sum benefit and proportional tax. The constrained optimal planner chooses:

$$b(j) = w(j)(1 - \tau)$$
$$\tau = \frac{\int_j n^{cm}(j)\theta(j)b(j)dj}{Y}$$

Here, individuals in each occupation consume $w(j)(1 - \tau)$ in each period, regardless of disability status. Then for $EU_j = EU_{j'}$, it must be $w(j) = w(j')$ for all j, j' and thus the constrained optimal planner achieves welfare maximization and output maximization simultaneously. $\qquad\square$

B.8. Extension: Heterogeneous Beta and Occupational Sorting

Here we expand the model to exhibit the kind of sorting on individual heterogeneity which we worked to control for in the empirical section. Specifically, we consider heterogeneity in the discount factor. We provide proofs for a population with two-types: patient types with high-beta β^H and impatient types with low-beta β^L. The measure of each type is fixed in each generation: $\phi^L \geq 0, \phi^H \geq 0$ such that $\phi^L + \phi^H = 1$. [15]

Our goal is to show the following. First, more patient individuals sort into higher risk occupations. Second, the main results of the paper go through: (i) the laissez-faire allocation is inefficient; and (ii) an economy with the marginal introduction of SDI from zero generates higher welfare for all agents compared to an economy with no SDI. We discuss interesting implications for unequal gains from SDI based on these results in the text.

---

[15]Our results do generalize to N types.

The Social Planner's Problem.. The allocation of labor in the social planner's problem is unchanged. The ratio of labor in two occupations are proportional to their risks.

Step 1: Separation of low and high-beta occupations. Want to show that the more patient high-beta types sort into the least risky occupations. To do this, we will show that if the high-beta type is indifferent between the risk-free occupation and any risky occupation, then the low-beta type strictly prefers the risky occupation. Thus, the low-beta type must inhabit strictly riskier occupations.

Proposition B.8 (Monotone Occupational Sorting). *There exists a unique $\bar{j}$ such that any occupation with $j < \bar{j}$ employs only high-beta types and any occupation with $j > \bar{j}$ employs only low-beta types.*

Proof. Observe that for arbitrary occupations $j < k$ that employ a high-beta type, it must be: $E_j[U^h] = E_k[U^h]$; and the same condition holds for low-beta types. Now, want to show for arbitrary occupations $j < k$ such that $E_j[U^h] = E_k[U^h]$, it must be $E[U^\ell(j)] < E[U^\ell(k)]$, which implies that occupation j will not employ a low-beta type.

WLOG assume $\beta^H = 1$ and denote β^L as just β. Let high-types be indifferent between occupations $j < k$.

$$u(w^*(k) - a^{*H}(k)) + \theta(k)u(a^{*H}(k)) + (1 - \theta(k))u(w^*(k) + a^{*H}(k)) = u(w^*(j) - a^{*H}(j)) + \theta(j)u(a^{*H}(j)) + (1 - \theta($$

$$= 2u(w^*(0))$$

Want to show:

$$u(w^*(j) - a^{*L}(j)) + \beta\theta(j)u(a^{*L}(j)) + \beta(1 - \theta(j))u(w^*(k) + a^{*L}(j)) > (1 + \beta)u(w^*(0))$$

Add $(1 - \beta)u(w^*(0))$ to both sides. Suffices to verify:

$$u(w^*(j) - a^{*L}(j)) + \beta\theta(j)u(a^{*L}(j)) + \beta(1 - \theta(j))u(w^*(j) + a^{*L}(j)) + (1 - \beta)u(w^*(0))$$

$$> (1 + \beta)u(w^*(0)) + (1 - \beta)u(w^*(0))$$

$$= u(w^*(j) - a^{*H}(j)) + \theta(j)u(a^{*H}(j)) + (1 - \theta(j))u(w^*(j) + a^{*H}(j))$$

20

By optimality of $a^{*L}(j)$, we have:

$$u(w^*(j) - a^{*L}(j)) + \beta\theta(j)u(a^{*L}(j)) + \beta(1 - \theta(j))u(w^*(j) + a^{*L}(j)) + (1 - \beta)u(w^*(0)))$$

$$\geq u(w^*(j) - a^{*H}(j)) + \beta\theta(j)u(a^{*H}(j)) + \beta(1 - \theta(j))u(w^*(j) + a^{*H}(j)) + (1 - \beta)u(w^*(0))$$

It then suffices to show

$$u(w^*(j) - a^{*H}(j)) + \beta\theta(j)u(a^{*H}(j)) + \beta(1 - \theta(j))u(w^*(j) + a^{*L}(j)) + (1 - \beta)u(w^*(0))$$

$$> u(w^*(j) - a^{*H}(j)) + \theta(j)u(a^{*H}(j)) + (1 - \theta(j))u(w^*(j) + a^{*H}(j))$$

Algebraically, this inequality holds if and only if $u(w^*(0)) > \theta(j)u(a^{*H}(j)) + (1 - \theta(j))u(w^*(j) + a^{*H}(j))$ But, by assumption,

$$2u(w(0)^*) = u(w^*(j) - a^{*H}(j)) + \theta(j)u(a^{*H}(j)) + (1 - \theta(j))u(w^*(j) + a^{*H}(j))$$

We have our result if $u(w^*(j) - a^{*H}(j)) > \theta(j)u(a^{*H}(j)) + (1 - \theta(j))u(w^*(j) + a^{*H}(j))$. Suppose not, then by the concavity of u it must be, $a^{*H}(j) > w^*(j) - a^{*H}(j) \Rightarrow a^{*H}(j) > \frac{1}{2}w^*(j)$. But this violates the first order condition:

$$u'(w^*(j) - a^{*H}(j)) \geq u'(\frac{1}{2}w^*(j))$$

$$\geq \theta(j)u'(\frac{1}{2}w^*(j)) + (1 - \theta(j))u'(3/2w^*(j))$$

$\square$

Step 2: CM with Heterogenous Beta is Pareto Inefficient. In this section we use the result from the previous proposition to argue that the competitive allocation is inefficient both in terms of consumption sharing and occupational sorting. In particular, we will show it is inefficient within beta-types, a simple extension of our proofs for the homogenous beta case.

Proposition B.9 (The Competitive Allocation with Heterogenous Beta and without Social Insurance is Pareto Inefficient). *Let $\{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}_{j=1}^J$ satisfy the definition of a competitive equilibrium in the case of (i) a continuum of occupations; (ii) two types of agents with different discount factors; (ii) and no social insurance ($b = \tau = 0$).*

There exists an alternative feasible allocation $\{\hat{c}_1(j), \hat{c}_d(j), \hat{c}_n(j), \hat{\ell}(j)\}_{j=1}^J$ that:
(i) increases expected utility in each occupation

$$E_j[U^m(\hat{c}_1(j), \hat{c}_n(j), \hat{c}_d(j))] \geq E_j[U^m(c_1^*(j), c_n^*(j), c_d^*(j))] \;\; \forall j \in \{1, \dots, J\}; \quad m \in \{L, H\}$$

$\exists k \;\; s.t. \;\; E_k[U(\hat{c}_1(k), \hat{c}_n(k), \hat{c}_d(k))] > E_k[U(c_1^*(k), c_n^*(k), c_d^*(k))]$

(ii) is feasible

$$\int_j \hat{\ell}(j) \left(\hat{c}_1(j) + \theta(j)\hat{c}_d(j) + (1 - \theta(j))\hat{c}_n(j)\right) dj \leq \left(\int_j (\hat{\ell}(j)(2 - \theta(j)))^\gamma dj\right)^{\frac{1}{\gamma}}$$

Proof. This is an application of Proposition 3.3. Fix $\bar{j}$ dictating the span of occupations of each type in the LF economy according to the above monotone sorting proposition. Next, open complete markets within each beta type and allow occupation choices within $j \in [\bar{j}, J)$ for the low-beta types and $j \in [0, \bar{j})$ for the high-beta types. By Proposition 3.3, the allocation will be preferred by each type and increase output. $\qquad\square$

Proposition B.10 (The Competitive Allocation with Heterogenous Beta and without Social Insurance is Puts too Few Workers in Risky Occupations). *Let $\{\ell_j^*\}$ satisfy the definition of a competitive equilibrium in the case of (i) a continuum of occupations; (ii) two types of agents with different discount factors; (ii) and no social insurance ($b = \tau = 0$). Let $\{\ell_j^{CM}\}$ be the feasible output maximizing allocation. Then:*

$$\int_{j=0}^t \ell_j^* \leq \int_{j=0}^t \ell_j^{CM} \quad \forall t \in [0, J]$$

ie: the efficient distribution of labor across occupations first-order stochastic dominates the distribution in the competitive equilibrium.

Proof. Fix $\bar{j}$ dictating the span of occupations of each type in the LF economy according to the above monotone sorting proposition. It is a simple application of Proposition 3.4 to show first order stochastic dominance within the occupations held by a given beta type, when normalized by the total number of workers in occupations either greater or less than $\bar{j}$.

$$\frac{\int_{j=0}^t \ell_j^*}{\int_{j=0}^{\bar{j}} \ell_j^*} \leq \frac{\int_{j=0}^t \ell_j^{CM}}{\int_{j=0}^{\bar{j}} \ell_j^{CM}} \quad \forall t \in [0, \bar{j}] \tag{B.35}$$

$$\frac{\int_{j=\bar{j}}^s \ell_j^*}{\int_{j=\bar{j}}^J \ell_j^*} \leq \frac{\int_{j=\bar{j}}^s \ell_j^{CM}}{\int_{j=\bar{j}}^J \ell_j^{CM}} \quad \forall s \in (\bar{j}, J] \tag{B.36}$$

22

What remains to show is that the LF economy puts more workers in the less risk occupations, those below $\bar{j}$, than the efficient allocation:

$$\phi^* \equiv \int_{j=0}^{\bar{j}} \ell_j^* \geq \int_{j=0}^{\bar{j}} \ell_j^{cm}$$

where ϕ^* equals the measure of high-beta type agents in each generation. Given B.35, it is sufficient to show $\ell^*(0) > \ell^{cm}(0)$ to show B.37 holds. This follows the same direct argument as the baseline case since, again, we are only dealing with relative quantities within a β type. $\qquad \square$

Step 3: Welfare Gain from SDI w/ Heterogenous Beta. The prior two steps have shown the competitive allocation of the economy with heterogenous beta generates the same sources of inefficiency, lack of risk sharing and productive inefficiency, as the economy with just one beta type. In this final step we show the marginal introduction of social disability insurance (SDI) continues to increase welfare and does so for both beta types.

Proposition B.11 (Social Insurance is Welfare Improving (on the margin) for all Beta–Types). *Let $EU^i(\tau)$ be the expected utility of a type $\beta^i \in \{\beta^h, \beta^\ell\}$ agent from the competitive equilibrium in the case of (i) a continuum of occupations; (ii) two types of agents with different discount factors; (ii) and proportional social insurance ($\tau \geq 0$, $b_j = bw_j$, and $\sum_j (2 - \theta_j)\ell_j \tau w_j = \sum_j \theta_j \ell_j bw_j$). Then:*

$$EU'^i(0) > 0 \quad \forall i \in \{\ell, h\}$$

Proof. **Lemma 1. SDI increases output ($\frac{dy^*}{d\tau}|_{\tau=0} > 0$).** Fix $\bar{j} \in (0, J]$. Impose a modified competitive economy with the additional restriction that high-beta types may only choose occupations $j \leq \bar{j}$ and low-beta types may only choose occupations $j > \bar{j}$. By an application of Lemma C.6, we have $\frac{dy^*}{d\tau}|_{\tau=0} > 0$, maintaining fixed $\bar{j}$. Denote $\bar{j}(\tau)$, as the high versus low beta cut-off in the economy with SDI characterized by $\tau \in [0, \epsilon]$ and $\bar{j}^{cm}$ as the cut-off in the first-best planner (output maximizing) economy. What remains to be shown is that:

$$\bar{j}(0) \leq \bar{j}(\tau) \leq \bar{j}^{cm}$$

Suppose this is not true, that at marginal occupation k such that occupation k employs high-beta types and occupation $k + \varepsilon$ employs low-beta types for all $\varepsilon > 0$, we have:

$E_j[U^H(w(j,0))] = E_k[U^H(w(k,0))]$ for all $j < k$, but $E_j[U^H(w(j,\tau))] > E_k[U^H(w(k,\tau))]$ for some $j < k$.

At $\tau = 0$, competitive wages must be monotone increasing, otherwise expected income would be higher in a less risky occupation and the occupation indifference condition cannot be satisfied. This implies that $w(k + \varepsilon, 0) > w(k, 0)$. Let $k - \delta < \bar{j}(\tau)$ The assumption for contradiction that $\bar{j}(0) > \bar{j}(\tau)$ combined with Lemma C.6 imply $n(k - \delta, \tau) > n(k - \delta, 0)$ for δ small, because the allocation under τ covers fewer occupations and puts more mass at riskier occupations. Subsequently: $w(k - \delta, \tau) < w(k - \delta, 0)$. Likewise, for occupations k, we have assumed $k > \bar{j}(\tau)$ and by the same logic applying C.6 $n(k, \tau) < n(k, 0)$ for δ small, because the allocation under τ covers more occupations and puts more mass at riskier occupations. Subsequently: $w(k, \tau) > w(k, 0) > w(k - \delta, 0) > w(k - \delta, \tau)$.

Let $E_j(y(w(j), \tau))$ be the total expected income in occupation j under SDI scheme τ. We then have:

$$E_k(y(w(k), \tau)) = (2 - \theta(k))w(k, \tau)(1 - \tau) + \theta(k)w(k, \tau)b(\tau) >$$
$$> E_k(y(w(k), 0)) = (2 - \theta(k))w(k, 0) >$$
$$E_{k-\delta}(y(w(k - \delta), 0)) = (2 - \theta(k - \delta))w(k - \delta, 0)$$

Where the first inequality is true for τ small since $b(\tau) > \tau$ at all values and $w(k, \tau) > w(k, 0) > w(k - \delta, 0)$. But, we have assumed $E_k[U^H(w(k, 0))] = E_{k-\delta}[U^H(w(k - \delta, 0))]$, then occupation k at tax rate τ represents a lottery with less risk and a strictly higher risk premium relative to occupation $k - \delta$ and it must be: $E_k[U^H(w(k, \tau))] > E_{k-\delta}[U^H(w(k - \delta, \tau))]$, a contradiction.

Part 2. SDI increases welfare for both types. Given the prior lemmas, proofs from the baseline case can be applied directly for high-beta types. What remains is to show low-beta types also gain. The thrust is similar to the baseline: so long as τ is sufficiently small that full insurance is not provided ($a^*(j, \tau) > 0$ for all j) and output increases, then we have a welfare improvement at τ for all types.

First, we argue that the introduction of disability insurance raises expected consumption of the workers in the highest risk occupation, $j = J$. Expected consumption in occupation

j for an arbitrary τ is:

$$C(j) \equiv w_j[(1-\tau)(2-\theta_j) + b\theta_j]$$
$$= n_j^{\gamma-1}y^{1-\gamma}[(1-\tau) + \frac{\theta}{2-\theta}b]$$

The second equality comes following some algebra after substitution of the competitive wage: $w_j = (\frac{y}{n_j})^{-1\gamma}$. The marginal change in total consumption at the introduction of SDI is:

$$\frac{dC(j)}{d\tau}\|_{\tau=0} = \frac{dn_j}{d\tau}(1-\gamma)(n_j)^{\gamma-2}y^{1-\gamma} + (1-\gamma)y^{-\gamma}n_j^{\gamma-1}\frac{dy}{d\tau} + n_j^{\gamma-1}y^{1-\gamma}[\frac{\theta_j}{2-\theta_j}\frac{db}{d\tau} - 1]$$
$$= \gamma w_j\frac{dn_j}{d\tau} + (1-\gamma)y^{-\gamma}n_j^{\gamma}\frac{dy}{d\tau} + n_j w_j[\frac{2}{2-\theta_j}\frac{db}{d\tau} - 1]$$

Recall: $b = \tau\frac{y}{\int_j \theta(j)w(j)\ell(j)dj}$. Since $\theta_j \in [0, J]$, the highest possible dependency ratio is one to three, and so we know $b \geq 3\tau$. Additionally, we have the prior results that (1) $n(\tau) \succ_{fsd} n(0)$ and so $\frac{dn(J)}{d\tau}|_{\tau=0} > 0$; and (2) $\frac{dy}{d\tau}|_{\tau=0}$. Together, these things imply that, given our parameter restriction $\gamma \in (-\infty, 1)$, we have $\frac{dC(j)}{d\tau}\|_{\tau=0} > 0$.

Now we argue $\frac{dE_J[C(J)]}{d\tau}\|_{\tau=0} > 0$ implies $d\frac{dE_JU(w(J,\tau))}{d\tau}|_{\tau=0} > 0$ which together with (1) the occupation indifference condition combined with (2) the sorting of low-beta types into the highest risk occupations; will imply $d\frac{dE_jU(w(j,\tau))}{d\tau}|_{\tau=0} > 0$ for all low-beta types (ie: occupations $j > \bar{j}(\tau = 0)$).

Let the optimal choice of assets in occupation $j = J$ at no SDI ($\tau = 0$) be denoted:

$$a^*(J, 0) \equiv \underset{a\geq 0}{\operatorname{argmax}}\, u((1-a)w(J,0)) + \beta\theta(J)u(aw(J,0)) + \beta(1-\theta(J))u((1+a)w(J,0))$$

The Inada condition $\lim_{c\to 0} u'(c) = \infty$ ensures that for all $\beta > 0$, and $N \in \mathbf{N}$, $\exists\ \ \delta > 0$ such that $\forall\ \ \tau \in (0, \delta)$ we have $a^*(J, 0) > \frac{N}{\beta}\tau$. Then, for N large, we can find a τ small such that $a^*(J, \tau) > 0$. At such a τ, Jensen's inequality provides $E_J[U(w(J,\tau))] > E_J[U(w(J,0))]$ because the former has higher expected payoff with lower variance.

$\square$

B.9. Extension: Costly Disability

Here we can further generalize the model to include a cost of becoming disabled. We will model it as a function $u^d(c)$ with the following properties: $u'^d(c) > u'(c)$ and $u^d(c) < u(c)$

for all $\infty > c > 0$. In words, this captures both an expenditure cost of disability through marginal utility and a utility cost through the absolute value. To simplify the analysis we assume further: $u'^d(c(1+\chi)) = u'(c)$ and $u^d(c(1+\chi)) = u(c)$ for all $\infty > c > 0$. That is the cost of disability is a constant proportion of consumption $(1+\chi)$ required to regain the utility of the non-disabled. In this way, we have a baseline where these costs have constant marginal relationships with the level of consumption and we can isolate factors in the planner's problem cleanly.

Step 1: Characterization of the First Best Allocation. The problem of a first-best social planner is similar to the base case except (i) relative consumption:

$$\max_{\ell(j),c_d(j),c_n(j),c_1(j)} \int_j [u(c_1(j)) + \theta(j)u^d(c_d(j)) + (1-\theta(j))u(c_n(j))]dj$$

$$st \int_j \ell(j)[c_1(j) + \theta(j)c_d(j) + (1-\theta(j))c_n(j)]dj \le \left(\int_j (2-\theta(j))^\gamma \ell(j)^\gamma dj\right)^{1/\gamma}$$

$$\int_j \ell(j)dj = 1$$

We see this problem is similar to the base case except: (i) relative consumption in the disabled and non-disabled states must be chosen; and (ii) the planner considers the differential marginal utility of consumption when allocating agents.

First order conditions straight-forwardly deliver:

$$u'(c_1) = u'(c_n) = u'^d(c_d) \quad \forall j$$

which implies: $c \equiv c_1 = c_n < c_d$.

Next we consider the allocation of labor across occupations. First order conditions imply:

$$w_j = c + \frac{\theta(j)}{2-\theta(j)}c_d$$

$$\frac{w_j}{w_0} = 1 + \frac{\theta(j)}{2-\theta(j)}\frac{c_d}{c}$$

$$\ell(j) = \frac{2}{2-\theta(j)}[1 + \frac{\theta(j)}{2-\theta(j)}\frac{c_d}{c}]^{\frac{1}{1-\gamma}}\ell(0)$$

$$= y(2-\theta_j)^{\frac{\gamma}{1-\gamma}}[(2-\theta_j)c + \theta_j c^d]^{\frac{1}{\gamma-1}}$$

$$\ell(0) = [\int_j (\frac{2}{2-\theta})(1 + \frac{\theta(j)}{2-\theta(j)}\frac{c_d}{c})^{\frac{1}{1-\gamma}}dj]$$

26

The last equation exploits the resource constraint on total labor per generation to be one.

Equivalence to the base case without a cost of disability is shown by setting $\chi = 0$. We see the disability cost drives a wedge between the relative marginal product of labor between two occupations of different risk. Increasing this cost χ increases the difference in the marginal product of labor between two occupations. The direct implication of this is that increasing these costs decreases relative labor in riskier occupations compared to safer occupations.

In a sense this is a second cost of disability in our model. The first cost was the "human capital" cost of fewer workers given by the fraction $\theta(j)$ who cannot work in the second period. The key difference is that the impact of the human capital cost on the planner's allocation depends on the elasticity of substitution across occupations. For $\gamma > 0$, the occupations are gross complements and the planner actually puts more labor in risky occupations relative to a safe ones. The qualitative impact of the utility costs do not depend on this elasticity, but the magnitude of the effect is increasing in the substitutability of occupations.

Step 1: Competitive Equilibrium with Utility Cost of Disability is Pareto Inefficient. In this section we use the result from the previous proposition to argue that the competitive allocation is inefficient both in terms of consumption sharing and occupational sorting.

Proposition B.12 (The Competitive Allocation with Utility Cost of Disability and without Social Insurance is Pareto Inefficient)**.** *Let* $\{c_{j,1}^*, c_{j,n}^*, c_{j,d}^*, a_j^*, n_j^*, \ell_j^*\}_{j=1}^J$ *satisfy the definition of a competitive equilibrium in the case of (i) a continuum of occupations; (ii) cost of disability; (ii) and no social insurance ($b = \tau = 0$). There exists an alternative feasible allocation* $\{\hat{c}_{j,1}, \hat{c}_{j,d}, \hat{c}_{j,n}, \hat{\ell}_j\}_{j=1}^J$ *that:*
(i) increases expected utility in each occupation

$$E[U_j(\hat{c}_{j,1}, \hat{c}_{j,n}, \hat{c}_{j,d})] \geq E[U_j(c_{j,1}^*, c_{j,n}^*, c_{j,d}^*)] \; \forall j \in \{1, \ldots, J\}$$
$$\exists k \; s.t. \; E[U_k(\hat{c}_{k,1}, \hat{c}_{k,n}, \hat{c}_{k,d})] > E[U_k(c_{k,1}^*, c_{k,n}^*, c_{k,d}^*)]$$

(ii) is feasible

$$\sum_j \hat{\ell}_j \left(\hat{c}_{j,1} + \theta_j \hat{c}_{j,d} + (1 - \theta_j)\hat{c}_{j,n}\right) \leq \left(\sum_j (\hat{\ell}_j (2 - \theta_j))^\gamma\right)^{\frac{1}{\gamma}}$$

27

Proof. This is obvious. We can construct a dominating allocation by fixing labor choices and allowing risk sharing within occupations as shown in the base case. Simply, for any occupation $j > 0$, we have $a^*(j) > 0$ in the competitive equilibrium with incomplete markets. This implies

$$E_j U(w^*(j)) = u(w^*(j) - a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j) + a^*(j))$$

Consider instead $\hat{c}_1(j) = w^*(j) - a^*(j) + (1 - \theta)a^*$; $\hat{c}_d(j) = a^*(j)$; and $\hat{c}_n(j) = w^*(j)$. This is feasible and gives:

$$E_j U(\hat{c}) = u(w^*(j) - \theta(j)a^*(j)) + \theta(j)u(a^*(j)) + (1 - \theta(j))u(w^*(j)) > E_j U(w^*(j))$$

$\square$

Step 2: Competitive Equilibrium with Utility Cost of Disability is Productively Inefficient.

Proposition B.13 (The Competitive Allocation with Utility Cost of Disability and without Social Insurance Does not Attain Productive Efficiency of the Complete Markets). *Let $\ell^*(j)$ satisfy the definition of a competitive equilibrium in the case of (i) a continuum of occupations; (ii) cost of disability; (ii) and no social insurance ($b = \tau = 0$). Let $\ell^{CM}(j)$ be the feasible output maximizing allocation. Then:*

$$y^* = \left(\int_{j=0}^{\bar{J}} (\ell^*(j))^\gamma dj\right)^{\frac{1}{\gamma}} < \left(\int_{j=0}^{\bar{J}} (\ell^{CM}(j)dj)^\gamma dj\right)^{\frac{1}{\gamma}} = y^{CM}$$

ie: the efficient distribution of labor across occupations first-order stochastic dominates the distribution in the competitive equilibrium.

Proof. This proof follows closely the analogous proof for the baseline model. Suppose, for contradiction that $\exists j > 0$ such that:

$$\frac{\ell(j)}{\ell(0)} \geq \frac{\ell^{CM}(j)}{\ell^{CM}(0)}$$

Using the definition of competitive wages, this would imply:

$$\frac{w(j)}{w(0)} \leq \frac{w^{CM}(j)}{w^{CM}(0)}$$

Since $u()$ is monotone and increasing, this implies:

$$2u(w(0)) > 2u(\frac{w^{CM}(0)}{w^{CM}(j)}w(j))$$

Let $c_d^{CM} = (1 + \chi)c^{CM}$. Our assumptions provide $\chi > 0$. With algebra on the social planner's optimality of ℓ (B.37), we find:

$$\frac{w_j^{CM}}{w_k^{CM}} = \frac{2 - \theta_k}{2 - \theta_j}\frac{2 + \chi\theta_j}{2 + \chi\theta_k}$$

and in the case of $j = 0$ and $k = j$:

$$\frac{w_0^{CM}}{w_j^{CM}} = \frac{2 - \theta_j}{2 + \chi\theta_j}$$

The occupational indifference condition in the competitive equilibrium requires:

$$\max_{a \geq 0} u(w(j) - a) + \theta(j)ud(a) + (1 - \theta(j))u(w(j) + a) = 2u(w(0))$$

Yet, we have just shown:

$$\max_{a \geq 0} u(w(j) - a) + \theta(j)ud(a) + (1 - \theta(j))u(w(j) + a) = 2u(w(0))$$

$$\geq 2u(\frac{2 - \theta_j}{2 + \chi\theta_j}w(j))$$

$$> 2u(\frac{2 - \theta_j}{2}w(j))$$

This means the incomplete markets, risky allocation, is strictly preferred to consuming half of expected earnings as a healthy person, in each period. This certainly violates Jensens' inequality and provides a contradiction.

$\square$

As in the baseline model without cost of disability, we have the corollary that $\frac{\ell(0)}{\ell^{CM}(0)} > 1$; the competitive allocation without insurance puts more workers in the risk-free occupation.

Step 3: Comparing the Competitive Allocation of Labor with Utility Cost of Disability to the Complete Markets Allocation..

Proposition B.14 (The Competitive Allocation with Utility Cost of Disability and without Social Insurance Puts too Few Workers in Risky Occupations). *Let $\ell^*(j)$ satisfy the definition of a competitive equilibrium in the case of (i) a continuum of occupations; (ii) cost of disability; (ii) and no social insurance ($b = \tau = 0$). Let $\ell^{cm}(j)$ be the feasible output maximizing allocation. Then:*

$$\int_{j=0}^{t} \ell^*(j)dj \leq \int_{j=0}^{t} \ell^{cm}(j)dj \quad \forall t \in [0, J]$$

ie: the efficient distribution of labor across occupations first-order stochastic dominates the distribution in the competitive equilibrium.

Proof. This proof follows closely the analogous proof for the baseline model, Proof B.4. Suppose, for contradiction that $\exists j < k$ such that:

$$\frac{\ell^*(k)}{\ell^*(j)} \geq \frac{\ell^{CM}(k)}{\ell^{CM}(j)}$$

Using the definition of competitive wages, this would imply:

$$\frac{w(k)}{w(j)} \leq \frac{w^{CM}(k)}{w^{CM}(j)}$$

and since $\frac{w_j^{CM}}{w_k^{CM}} = \frac{2-\theta_k}{2-\theta_j}\frac{2+\chi\theta_j}{2+\chi\theta_k}$:

$$w^*(k)(2 - \theta(k))(2 + \theta(j)\chi) < w^*(j)(2 - \theta(j))(2 + \theta(k)\chi)$$

Since $\chi > 0$ and we are given $\theta(j) < \theta(k)$, the result is that the total expected income in occupation k is less than in occupation j

$$w^*(k)(2 - \theta(k)) < w^*(j)(2 - \theta(j))$$

We have reduced the problem to the baseline case. Together we have lower expected income and higher risk in occupation k. This arrangement cannot satisfy indifference in expected utilities across occupations. See proof for the baseline case for a formal exposition. $\square$

30

Step 4: Welfare Gain from SDI w/ Costly Disability. The prior two steps have shown the competitive allocation of the economy with costly disability generates the same sources of inefficiency, lack of risk sharing and productive inefficiency, as the baseline economy. In this final step we show the marginal introduction of social disability insurance (SDI) continues to increase welfare.

Proposition B.15 (Social Insurance is Welfare Improving (on the margin)). *Let $EU(\tau)$ be the expected utility of an agent from the competitive equilibrium in the case of (i) a continuum of occupations; (ii) costly disability; and (iii) and proportional social insurance ($\tau \geq 0$, $b_j = bw_j$, and $\sum_j (2 - \theta_j)\ell_j \tau w_j = \sum_j \theta_j \ell_j bw_j$). Then:*

$$\frac{dEU(\tau)}{d\tau}\bigg\|_{\tau=0} > 0$$

Proof. **Lemma 1. SDI increases output ($\frac{dy^*}{d\tau}|_{\tau=0} > 0$).** First, we will show:

$$\frac{d}{d\tau}\left(\frac{\ell^*(j,\tau)}{\ell^*(0,\tau)}\right) > 0 \quad \forall j > 0$$

Following the analogous proof for the baseline case, we consider δ such that $\forall \epsilon \in (0,\delta)$, we have $a^*(j,\delta) > 0$ for every $j > 0$. That is to say, the borrowing constraint does not bind and choices are all in the interior of the budget set.

Hold wages constant at the no-insurance competitive level. Then for $j > 0$ the expected income at $\tau = 0$ and $\tau = \epsilon > 0$ are $(2 - \theta(j))w(j,0)$ and $(2 - \theta(j))w(j,0)(1 - \epsilon) + \theta(j)\hat{b}(\epsilon|0)w(j,0)$, respectively. We define $b(\epsilon)$ as the level of benefits implied by $\tau = \epsilon$ and $w(j,\epsilon)$.

The expected income at $\tau = \epsilon > 0$ translates to higher expected consumption under our assumption $u'(c) = u'^d((1+\chi)c)$ for all c, if $\frac{b(\epsilon)}{1+\chi} > \epsilon$. Since we restrict $\theta(j) \leq \frac{1}{2}$ we have at most $\frac{1}{4}$ of the population in each occupation disabled. Considering repayments only within occupation we would have a dependency ratio of 1-to-3 implying $b \geq 3\tau$. Therefore, a sufficient (although not necessary) condition for the remainder of our proofs to go through is:

$$(1+\chi) < 3 \Rightarrow \frac{3\tau}{1+\chi} > \tau \rightarrow \frac{b(\tau)}{1+\chi} > \tau$$

With this result, we apply the Maximum Theorem as in the baseline model to show expected utility is higher:

$$\max_{a_\epsilon} u(w(j,0)(1-\epsilon) - a_\epsilon) + \theta(j)u^d(w(j,0)b(\epsilon) + a_\epsilon) + (1-\theta(j))u(w(j,0)(1-\epsilon) + a_\epsilon)$$
$$> \max_{a_0} u(w(j,0) - a_0) + \theta(j)u^d(a_0) + (1-\theta(j))u(w(j,0) + a_0)$$

For $j = 0$, however, the expected utility is lower with the ϵ tax rate:

$$2u(w(0,0)(1-\epsilon)) < 2u(w(0,0))$$

Based on indifference under $\tau = 0$, we have

$$\max_{a_\epsilon} u(w(j,0)(1-\epsilon) - a_\epsilon) + \theta(j)u^d(w(j,0)b(\epsilon) + a_\epsilon) + (1-\theta(j))u(w(j,0)(1-\epsilon) + a_\epsilon)$$
$$> 2u(w(0,0)(1-\epsilon)) \tag{B.37}$$

This violates the indifference condition of the competitive equilibrium with $\tau = \epsilon$ and so we must have a change in both the wage in the $j = 0$ occupation and the relative wage in the $j > 0$ occupations: $w(j,\epsilon) \neq w(j,0)$ and $\frac{w(j,\epsilon)}{w(0,\epsilon)} \neq \frac{w(j,0)}{w(0,0)}$

The remainder of the proof showing it must be $w(j,\epsilon) < w(j,0)$ and $\frac{w(j,\epsilon)}{w(0,\epsilon)} < \frac{w(j,0)}{w(0,0)}$ follows directly the proof for the baseline case and so is omitted.

Corollary 1. Marginal SDI decreases labor in the risk free occupation: $\ell^*(0,\epsilon) < \ell^*(0,0)$ This follows a direct application of Corollary E.6, the analogous corollary for the baseline case.

Lemma 2. Output is increasing at the marginal introduction of SDI: $\frac{dy^*(\tau)}{d\tau}|_{\tau=0} > 0$ This follows a direct application of Lemma E.7, the analogous lemma for the baseline case.

Part 2. SDI increases welfare: $\frac{dEU(j,\tau)}{d\tau}|_{\tau=0} > 0$ **for all** j
This proof again follows from the baseline case once adjusting disability benefits to the non-disabled equivalent consumption: $\hat{b}(\tau) = \frac{b(\tau)}{1+\chi}$ and proceeding from there.

$\square$

B.10. Extension: Non-verifiable disability status

Proposition B.16 (The Optimal SDI Policy is Robust to Non-Verifiable Disability Status).
Let $\{c_1^*(j,\tau^*), c_n^*(j,\tau^*), c_d^*(j,\tau^*), a^*(j,\tau^*), w(j,\tau^*)\}_{j=1}^J$ *satisfy the definition of a competitive equilibrium at the optimal SDI policy* $(\tau, b = b(\tau))$. *Then, the expected utility of reporting*

disability status truthfully, $EU(j, \tau)$, is greater than any time-consistent deviation:

$$EU(j, \tau) = \max_a u((1-\tau)w(j, \tau) - a) + \theta(j)u(bw(j, \tau) + a) + (1 - \theta(j))u((1-\tau)w(j, \tau) + a)$$

$$\geq \max_a u((1-\tau)w(j, \tau) - a) + \theta(j)u(bw(j, \tau) + a)$$

$$+ (1 - \theta(j))\max\{u((1-\tau)w(j, \tau) + a), u(bw(j, \tau) + a)\}$$

Proof. We first show that a worker in any occupation $j \in [0, J]$ who enters the second period with an arbitrary asset $a > 0$ optimizes by truthfully revealing their disability status when healthy, ie:

$$u(bw(j, \tau) + a) \leq u(w(j, \tau)(1 - \tau) + a)$$

where the left-hand side is the payoff if the worker claims disability and right-hand side is the payoff from working. This holds if $b \leq 1 - \tau$, which we have shown in Lemma B.4 to be the case given our restriction $J \leq \frac{1}{2}$. We can quickly recap that result here: From the government budget constraint, given a τ, b satisfies

$$b \int_k \theta(k)w(k, \tau)dk = \tau \int_k (2 - \theta(k))w(k, \tau)dk$$

and therefore $b = \tau \frac{\int_k (2-\theta(k))w(k,\tau)dk}{\int_k \theta(k)w(k,\tau)dk}$. $\theta(k) < 1$ and $w(k, \tau) > 0 \; \forall k$, $\frac{\int_k (2-\theta(k))w(k,\tau)dk}{\int_k \theta(k)w(k,\tau)dk} < 1$ and thus $b \geq \tau$ and equivalently $1 - b \leq 1 - \tau$.

Finally, if $b < \frac{1}{2}$ then $b \leq 1 - b$ and $1 - b \leq 1 - \tau \Rightarrow b \leq 1 - \tau$.

Therefore, any time consistent plan requires truth-telling in the final period and so the agent cannot commit to lie from period 1. $\square$

	(1)	(2)	(3)	(4)
O*NET Phys Reqs	0.0350 **	0.0379 **	0.0325 **	0.0356 **
	(0.0099)	(0.0103)	(0.0098)	(0.0107)
Potential Experience	0.0327 **	0.0322 **	0.0327 **	0.0322 **
	(0.0013)	(0.0010)	(0.0013)	(0.0010)
BMI	0.0121 **	0.0146 **	0.0121 **	0.0146 **
	(0.0010)	(0.0011)	(0.0009)	(0.0011)
Woman	0.0322 **	0.0506 **	0.0312 **	0.0500 **
	(0.0107)	(0.0112)	(0.0106)	(0.0115)
Self-employed	-0.0216 †	-0.0140	-0.0218 †	-0.0142
	(0.0128)	(0.0130)	(0.0128)	(0.0135)
College	0.1483 **	0.1596 **	0.1473 **	0.1585 **
	(0.0132)	(0.0138)	(0.0128)	(0.0139)
No HS	-0.0921 **	-0.0842 **	-0.0912 **	-0.0835 **
	(0.0107)	(0.0116)	(0.0106)	(0.0121)
Not married	0.0850 **	0.0944 **	0.0851 **	0.0945 **
	(0.0093)	(0.0110)	(0.0093)	(0.0115)
Not white	0.0332 **	0.0445 **	0.0337 **	0.0449 *
	(0.0121)	(0.0142)	(0.0120)	(0.0147)
Smoker	0.0203 †	0.0244 †	0.0207 †	0.0246 †
	(0.0113)	(0.0129)	(0.0112)	(0.0134)
Observations	14558	14558	14558	14558
J statistic	0.4749	.		
p-value, J	0.7886			

Robust standard errors in parentheses

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 5: The effect of an occupation's physical requirements on working-life disability including industry fixed effects and clustering standard errors on industry. Columns (1) and (2) use our instrumental variable schemes and Columns (3) and (4) treat physical requirements as exogenous. (1) and (3) are marginal effects at the mean from probit models. (2) and (4) are coefficients of linear probability models.

O*NET Non-phys Reqs, PC1	-0.711 **	(0.00212)
O*NET Non-phys Reqs, PC2	0.269 **	(0.00336)
O*NET Non-phys Reqs, PC3	0.301 **	(0.00261)
Constant	-0.00922 **	(0.00251)
Observations	14558	
R^2	0.933	
F	104812.0	

Robust standard errors in parentheses

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 6: First stage estimates using non-physical descriptors to instrument physical requirements. We include industry fixed effects and clustering standard errors on industry

	(1)	(2)
Safe	0.152	0.069
Risky	0.206	0.094
Difference	-0.053	-0.025
Observables	-0.034	-0.017
% Difference	64.2	68.0
Occupation	-0.020	-0.008
% Difference	35.8	32.0
N	20,328	127,298

Table 7: The decomposition between higher and lower disability risk, measured by self-reported work limitations due to health. Column (1) uses only one observation per individual and (2) pools all of the data.

	(1)	(2)	(3)	(4)
O*NET Phys Reqs	0.0130 **	0.0129 **	0.0121 **	0.0122 **
	(3.41)	(3.28)	(3.36)	(3.22)
Potential Experience	0.0051 **	0.0051 **	0.0051 **	0.0051 **
	(7.80)	(7.78)	(7.81)	(7.79)
BMI	0.0029 **	0.0030 **	0.0030 **	0.0030 **
	(4.70)	(4.37)	(4.71)	(4.38)
Woman	-0.0026	-0.0020	-0.0031	-0.0023
	(-0.39)	(-0.31)	(-0.47)	(-0.35)
Self-employed	-0.0290 **	-0.0272 **	-0.0291 **	-0.0273 **
	(-3.81)	(-3.97)	(-3.82)	(-3.98)
College	-0.0331 **	-0.0268 **	-0.0336 **	-0.0271 **
	(-4.06)	(-3.54)	(-4.12)	(-3.60)
No HS	0.0448 **	0.0591 **	0.0452 **	0.0594 **
	(5.40)	(5.95)	(5.45)	(5.98)
Not married	0.0155 †	0.0162 †	0.0155 †	0.0162 †
	(1.77)	(1.72)	(1.77)	(1.73)
Not white	-0.0188 *	-0.0180 *	-0.0186 *	-0.0178 *
	(-2.24)	(-2.07)	(-2.21)	(-2.06)
Smoker	0.0198 **	0.0197 *	0.0199 **	0.0198 *
	(2.71)	(2.53)	(2.72)	(2.54)
Observations	14763	14763	14763	14763
J statistic	9.0786	2.2145		
p-value, J	0.0107	0.3305		

t statistics in parentheses

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 8: The effect of an occupation's physical requirements on working-life disability, measured by self-reported work limitations due to health.