Journal of Monetary Economics 96 (2018) 77–92

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jmoneco

Occupational hazards and social disability insurance[☆]

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

[a] Department of Economics, Indiana University, 100 S. Woodlawn Ave, Bloomington, IN 47405-7104, United States
[b] Stony Brook University, United States

ARTICLE INFO

Article history:
Received 25 August 2014
Revised 13 February 2018
Accepted 8 April 2018
Available online 10 April 2018

JEL classification:
E62
I13

Keywords:
Disability insurance
Occupational choice
Optimal policy

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 6.3% relative to actuarially fair insurance, mostly due to risk sharing.

1. Introduction

How do workers' occupational choices affect the optimality of social—or public—disability insurance (SDI)? The two are related because, in choosing their occupation, workers are choosing their physical health risk and, hence, their likelihood to need disability insurance. Disability insurance then alters the composition of occupations by incentivizing workers to chose occupations with higher risk. This reallocation towards risky occupations can, however, be desirable when occupations are imperfect substitutes in production and insurance markets are incomplete.

Our central results are based on a theoretical model of occupation choice, occupation-specific disability risk, and incomplete markets for private insurance. We show that fewer workers choose occupations with high disability risk than would achieve productive efficiency. This occurs because workers in high-risk jobs demand a compensating differential, a wage sufficient to self-insure against disability. But with a downward sloping demand for additional workers in each occupation, these high wages obtain only if there are inefficiently few workers in high-risk occupations. In other words, self-insurance is too expensive and so equilibrium prices correspond to an inefficient distribution of workers.

The introduction of SDI improves welfare through two channels. The first is through risk sharing: social insurance improves welfare by helping workers in risky occupations to smooth consumption. The second, novel channel is occupational reallocation: SDI encourages more workers to choose risky occupations. On the margin, the allocation of workers across occupations becomes more efficient and output increases. The gains are larger than the cost to fund the scheme because social

☆ We thank Victor Rios-Rull, Juan-Carlos Hatchondo, Jonathan Heathcote, John Kennan, Sam Schulhofer-Wohl, and Ali Shorideh; workshop participants at University of Illinois and University of California-Santa Barbara, and seminars at the Bureau of Labor Statistics, Notre Dame, Iowa, Washington University in St. Louis, and Windsor. James Eubanks provided excellent research assistance.

* Corresponding author.
E-mail address: ammichau@indiana.edu (A. Michaud).

https://doi.org/10.1016/j.jmoneco.2018.04.002
0304-3932/

insurance is more efficient than self-insurance. With self-insurance, some workers who accumulate savings never become disabled. The increase in output provides welfare gains even for workers with zero disability risk and many whose disability risk never realizes.

The two channels, risk-sharing and the reallocation, both work to increase welfare upon introduction of social insurance, but at the welfare-maximizing level they are at odds. Although reallocation towards risky occupations initially improves productive efficiency, additional insurance can increase the risky occupational allocation beyond the output maximizing level. At the point at which an SDI expansion reduces output, welfare gains from risk-sharing remain. This result is general and not a quantitative statement; welfare is maximized by an SDI program that induces workers to choose risky occupations beyond the output maximizing level.

Our results are qualitatively robust to ex ante heterogeneity in preferences that induce sorting and to private information over disability status. Further, private information over disability status 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. This adverse selection leads to a market failure typical of insurance contracts with private information.[1] Social insurance can dictate equal insurance coverage so it is not subject to adverse selection and always generates welfare gains.

To motivate our theoretical work, we show heterogeneity in disability incidence across occupations. Using data from the University of Michigan's Health and Retirement Survey Health and Retirement Study HRS, we find a natural grouping between low- and high-risk occupations, the latter have about twice the disability rate as the low-risk group. While this is suggestive, occupational choice is endogenous and potentially influenced by unobservable factors. To begin addressing this, we propose an instrument scheme using O*NET measures of physical and non-physical occupational requirements. Physical requirements have health repercussions, but their estimated effect may be influenced by sorting along these physical requirements. Therefore, we instrument physical requirements by the non-physical requirements in the O*NET. The intent is to use how occupations bundle requirements: While workers' unobserved physical traits may guide sorting along the physical requirements, they hopefully do not guide sorting on non-physical requirements, but, these requirements can predict the physical requirements. This instrumenting scheme is not a definitive solution, but it does uncover significant bias in the estimate assuming exogeneity.

We use these facts, along with the U.S. SSDI system, to calibrate our model of occupation risk. We find the optimal SDI program costs 4.9% of GDP and provides welfare gains equivalent to a 6.3% increase in consumption in a world with actuarially fair insurance alone. Relative to this optimal program, the current U.S. system captures 84% 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. In both, 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.

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.

Several papers discuss disability's interaction with other economic factors. Notably, Golosov and Tsyvinski (2006) also consider optimal DI, but from a mechanism design approach to prevent misreporting. Private information is a potentially important consideration, and so we show that workers reveal their status in our baseline SDI scheme because our policy tool is less generous than would tempt misreporting.

We motivate our normative work by estimating the heterogeneity in disability risk across occupations. A few papers present similar empirical results (e.g. Fletcher et al., 2011; Morefield et al., 2012). 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

[1] 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.

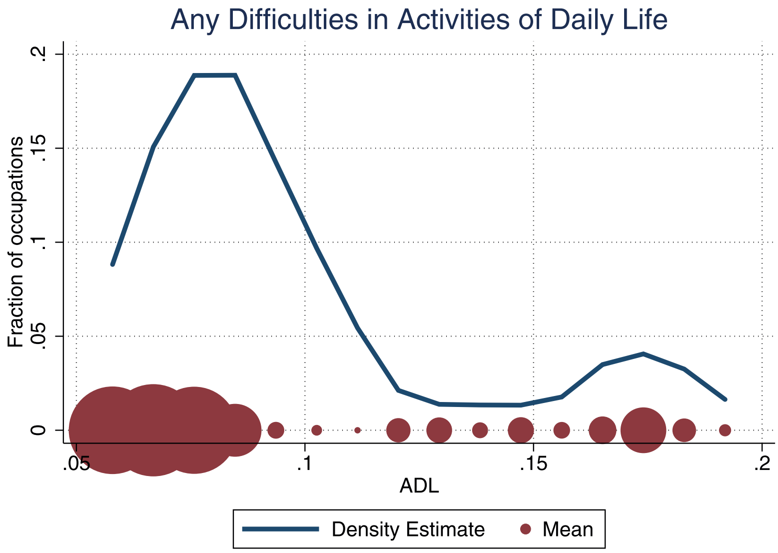

Fig. 1. The density across occupations of the incidence of difficulties with ADLs.

(HRS). The survey respondents describe their health conditions, which we compile into a measure of disability. This data shows that occupations have quite disparate disability rates and further analysis suggests much of it can be attributed to a occupations effects.

2.1. Constructing the dataset

RAND's distribution of the HRS forms the basis of our dataset, with further adjustments described in Online Appendix A.1. In it, there are several indicators of health problems during working life, before 65. Our baseline uses a relatively restrictive measure, tallying whether a respondent reports limitations to activities of daily living (ADLs). There is another, broader measure using a survey question about whether a health problem limits one's ability to do paid work. Results are generally very similar between the two, as shown in Table 2.

To get an occupation's health risk, individuals are associated with his or her longest-held occupation, reported retrospectively. Using current occupation instead would mistakenly consider an occupation dangerous if workers in poor health switched into that occupation. Each occupation is characterized by Knowledge, Skills and Abilities descriptors from O*NET version 4.0, "analyst," database. These are summarized by the first principal component from the 19 physical descriptors and three components from the 101 non-physical descriptors.

2.2. The distribution of disability across occupations

Fig. 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.

Observationally, there are two clusters of occupations, a larger low-incidence group and a tail 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 low-risk group. About 17% of the population worked in this high risk group.

The largest high-risk occupations are machine and transport operators and also include construction and extraction occupations. The large, low-risk occupations are professional and management occupations, administrative support and sales.

2.3. Demonstrative evidence of 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 individuals who choose what appear to be "risky" occupations and those who choose safe occupations. We first use an Oaxaca–Blinder decomposition to extract the residual effect of occupation on disability from observable differences between workers in different occupations. However, it is problematic to interpret this decomposition because the grouping is endogenous, i.e. people choose their occupations. We attempt to address the issue

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.

	(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 2
The effect of an occupation's physical requirements on working-life disability. Columns (1)–(3) use our instrumental variable schemes and Columns (4)–(6) treat physical requirements as exogenous. (1) and (3) are marginal effects at the mean from probit models. (3) and (6) use a self-reported health limitation as the dependent variable rather than an ADL difficulty.

	(1)	(2)	(3)	(4)	(5)	(6)
O*NET Phys	0.034 **	0.035 **	0.013 **	0.031 **	0.033 **	0.012 **
	(0.005)	(0.010)	(0.004)	(0.005)	(0.010)	(0.004)
Experience	0.033 **	0.033 **	0.005 **	0.033 **	0.033 **	0.005 **
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
BMI	0.012 **	0.012 **	0.003 **	0.012 **	0.012 **	0.003 **
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Woman	0.027 **	0.032 **	−0.003	0.026 **	0.031 **	−0.003
	(0.009)	(0.011)	(0.007)	(0.008)	(0.011)	(0.007)
Self-employed	−0.022 *	−0.022 †	−0.029 **	−0.022 *	−0.022 †	−0.029 **
	(0.010)	(0.013)	(0.008)	(0.010)	(0.013)	(0.008)
College	0.144 **	0.148 **	−0.033 **	0.142 **	0.147 **	−0.034 **
	(0.010)	(0.013)	(0.008)	(0.010)	(0.013)	(0.008)
No HS	−0.093 **	−0.092 **	0.045 **	−0.092 **	−0.091 **	0.045 **
	(0.011)	(0.011)	(0.008)	(0.011)	(0.011)	(0.008)
Not married	0.087 **	0.085 **	0.015 †	0.087 **	0.085 **	0.016 †
	(0.011)	(0.009)	(0.009)	(0.011)	(0.009)	(0.009)
Not white	0.031 **	0.033 **	−0.019 *	0.031 **	0.034 **	−0.019 *
	(0.010)	(0.012)	(0.008)	(0.010)	(0.012)	(0.008)
Smoker	0.021 *	0.020 †	0.020 **	0.021 *	0.021 †	0.020 **
	(0.009)	(0.011)	(0.007)	(0.009)	(0.011)	(0.007)
Industry FE		X			X	
Observations	14,763	14,558	14,763	14,763	14,558	14,763

Standard errors in parentheses † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

of endogenous occupation selection by introducing an instrument: We use O*NET measures of physical demands to summarize the occupation's effect on health, which we instrument by the non-physical demands that tend to be bundled together within occupations.

The Oaxaca–Blinder decomposition splits the differences in disability outcomes between observable characteristics and residual differences in disability rates across occupation groups. We group occupations into two sets, low- and high-risk, based on the mixture we estimated on Fig. 1, assigning each occupation to the distribution with highest density at its level of risk. Then we regress an indicator of whether the individual ever reported disability on the following regressors: a cubic for potential experience, body mass index (BMI), time, dummies for education level, gender, marital status, race, and tobacco use. We set the characteristics of workers who report a disability by their values in the wave in which they become disabled. For workers who never report a disability, we use the mean value of their characteristics across waves. The results of the decomposition are summarized in Table 1.

Column (1) of Table 1 averages covariates to make one observation per individuals; column (2) use multiple periods per individual. In our baseline case, Column (1), the occupation effect accounts for 42% of the difference in risk. This decomposition suggests much of the difference in risk is unexplained by population differences without differences in occupations themselves.

While this decomposition is instructive, there may also be unobservable characteristics that differ across workers in risky and safe occupations, and this implies there also may be sorting. For example, if more physically robust workers sort into physically-demanding occupations, this would attenuate occupations' effect. We now use instruments to partly address this.

We map occupations' health risk into the physical demands characterized by the O*NET, the construction of which we described in Section 2.1. To relate to our two occupation groups, risky occupations have a mean physical requirement of

1.05 and safe occupations have a mean of −0.65, where the population is mean 0 and standard deviation one. Using a direct measure of physical requirement instead of the identity of the occupation cuts away other characteristics of occupations that may be related to disability, such as economic decline. However, it 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. These requirements will be valid instruments if non-physical requirements are 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

$$\Pr[\text{Disabled}_i] = f(\gamma H_j + x_i'\beta + \nu_i), \tag{1}$$

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

The marginal effects from the estimation of Eq. 1 are in Table 2. Column (1) displays the baseline IV-Probit estimates. Column (2) adds industry fixed effects and clusters standard errors on industry, which may be important to control for economic factors affecting labor-force attachment and disability reporting. Column (3) uses health-limitations rather than ADL difficulties as the measure of disability. Columns (4)–(6) repeat these specifications but without instrumenting. The first stage of the IV, in Online Appendix Section A.2 which shows that non-physical requirements are strongly negatively related to physical requirements. To depart from distributional assumptions of the Probit, we also include estimates of the linear probability model and IV linear probability model in Online Appendix A.3. The coefficients are quantitatively very similar to the marginal effects of our Probit.

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. To the extent that our instrument is effectual, this suggests that people sort into more physically demanding occupations if they have a lower disability risk.

The results of our instrumental scheme are suggestive of the underlying relationship on which our theory is based; however, one could reasonably question the validity of these instruments. Essentially, if workers sort on comparative advantage or if their traits are negatively correlated then sorting on physical robustness also implies sorting on non-physical acuity, then the instrument will be invalid. Further, the actual task and skill composition of occupations are endogenous to a degree: the bundle is chosen in part as a result of the sorting of workers. Quantitatively, the first-stage is worryingly predictive—its F-statistic's p-value is 0.000—suggesting our instruments are too tightly related to the endogenous variable. Here, we have not definitively established a causal link but offered evidence of a relationship.

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:[3]

[2] 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.

[3] 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.

$$\max_{\{n(j)\}_{j=0}^{J}} y - \int_{j=0}^{J} w(j)n(j)dj \text{ s.t.} \tag{2}$$
$$y = \left(\int_{j=0}^{J} n(j)^{\gamma}\right)^{1/\gamma}$$

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 $(1-\tau)w(j)$. 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, $\theta(j)$.

$$E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))] = \max_{c_1, c_n, c_d, a} u(c_1) + (1-\theta(j))u(c_n) + \theta(j)u(c_d) \tag{3}$$
$$\begin{aligned} \text{s.t. } c_1 &\le w(j)(1-\tau) - a \\ c_n &\le w(j)(1-\tau) + a \\ c_d &\le a + bw(j) \\ a &\ge 0 \end{aligned}$$

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: $n^*(j) \le (2-\theta(j))\ell^*(j)$; 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: $w(j) = (\frac{y}{n(j)})^{1-\gamma}$
- 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

$$y \ge \int_{j=0}^{J} c_1(j) + (1-\theta(j))c_n(j) + \theta(j)c_d(j)dj \tag{4}$$

Government budget balance requires

$$b^* \le \frac{\tau y}{\int_j \theta(j)\ell(j)^* w(j)^* dj} \tag{5}$$

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. It solves*

$$\max_{\{c_1(j), c_d(j), c_n(j), \ell(j)\}} \int_j \ell(j)(u(c_1(j)) + \theta(j)u(c_d(j)) + (1-\theta(j))u(c_n(j)))dj \tag{6}$$
$$\text{s.t. } \left(\int_j ((2-\theta(j))\ell(j))^{\gamma} dj\right)^{\frac{1}{\gamma}} \ge \int_j \ell(j)(c_1(j) + \theta(j)c_d(j) + (1-\theta(j))c_n(j))dj$$
$$1 \ge \int_j \ell(j)dj$$

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 Online Appendix B.4. □

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. 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* $b = \tau = 0$. *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

$$E_j[U(\hat{c}_1(j), \hat{c}_n(j), \hat{c}_d(j))] \geq E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))] \ \forall j \in [0, J] \tag{7}$$
$$\exists k \text{ s.t. } E_k[U(\hat{c}_1(k), \hat{c}_n(k), \hat{c}_d(k))] > E_k[U(c_1^*(k), c_n^*(k), c_d^*(k))]$$

(ii) is feasible

$$\int_j \hat{\ell}(j)\big(\hat{c}_1(j) + \theta(j)\hat{c}_d(j) + (1-\theta(j))\hat{c}_n(j)\big)dj \leq \left(\int_j (\hat{\ell}(j)(2-\theta(j)))^{\gamma} dj\right)^{\frac{1}{\gamma}} \tag{8}$$

Proof. See Online Appendix B.2. □

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,*

$$\int_{j=0}^{t} \ell^*(j)dj \leq \int_{j=0}^{t} \ell^{cm}(j)dj \quad \forall t \in (0, J] \tag{9}$$

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

Proof. See Online Appendix B.2. □

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. Fig. 2 illustrates this for constant relative risk aversion (CRRA) preferences, $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, and $\gamma = 0.5$.

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 equivalent with all of the workers in the safest occupation. When production is Leontief, both allocations evenly distribute workers across occupations. Fig. 3 illustrates this, fixing risk aversion at $\sigma = 2$.

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 (There exists a strictly positive welfare maximizing level of insurance that is a Pareto improvement over no insurance.)**.** *Let EU(j, τ) be the expected utility* $E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))](\tau)$ *from the allocation in competitive equilibrium*

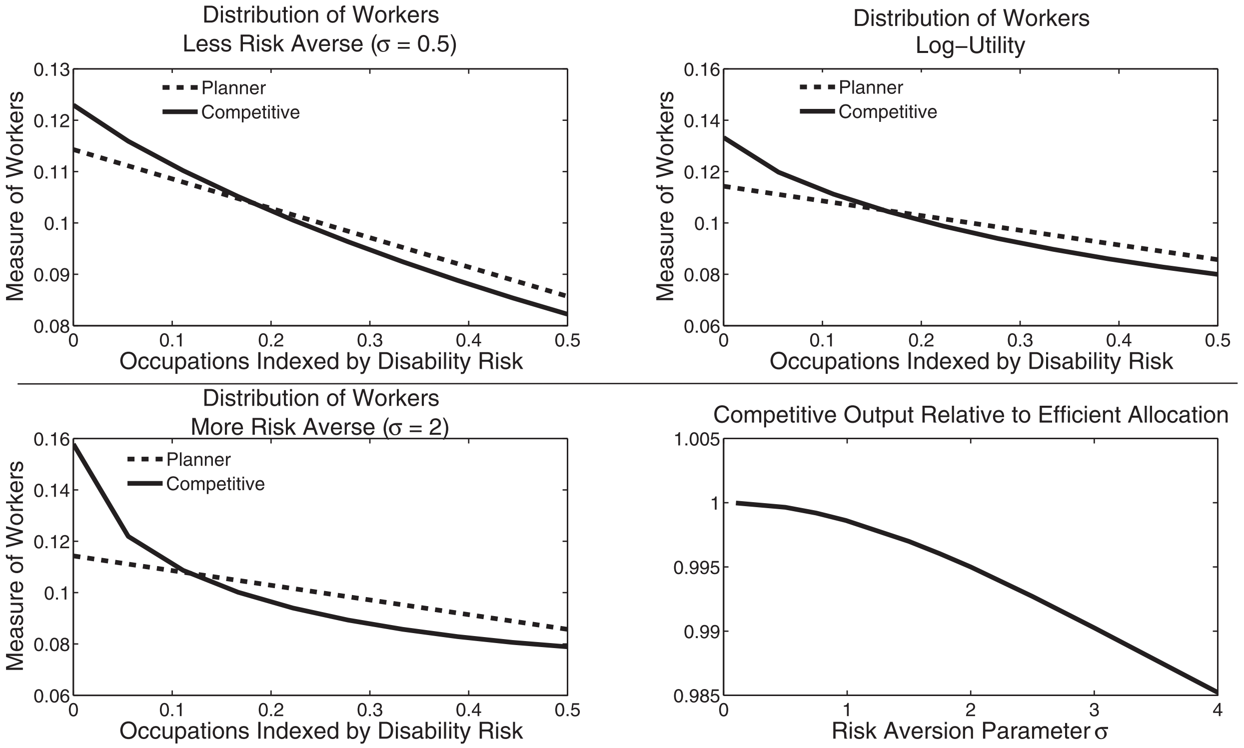

Fig. 2. The competitive allocation relative to the efficient, planner's at various levels of risk aversion. The labor allocation diverges more greatly from the optimal planner's at higher levels of risk aversion.

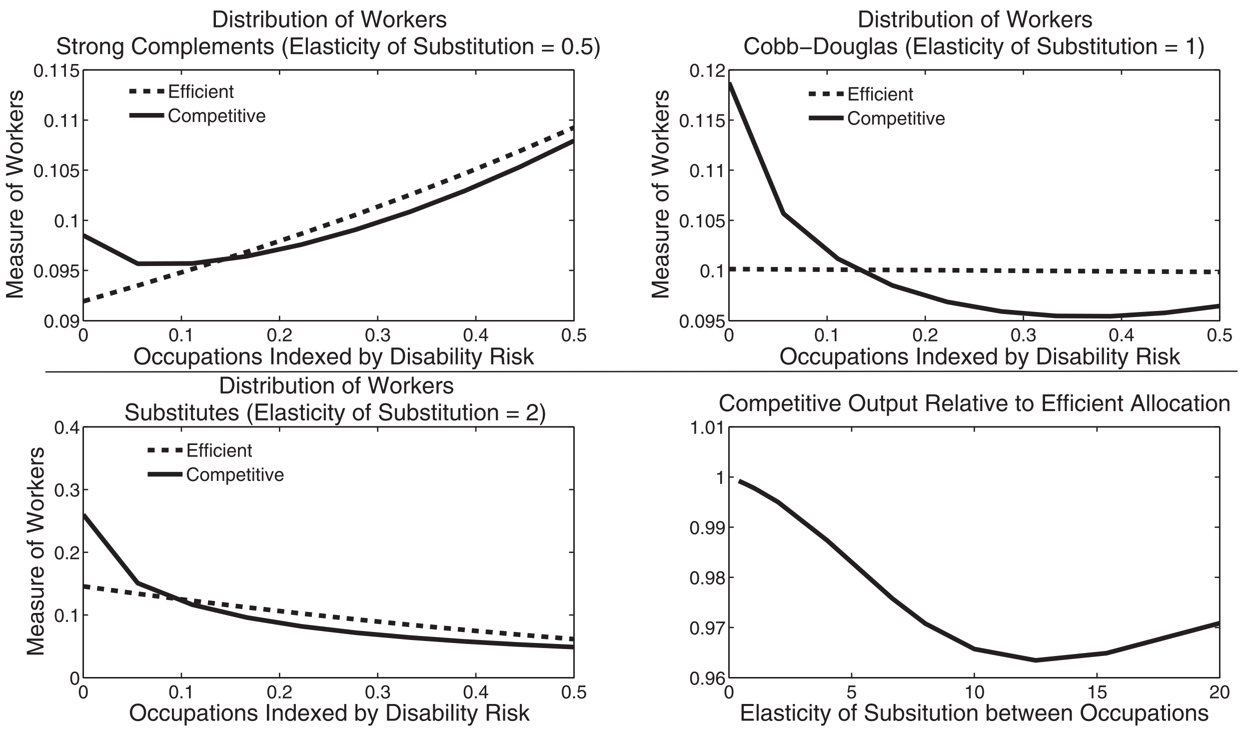

Fig. 3. The competitive allocations relative to the efficient at various elasticities of substitution. Distortions away from the optimal planner's increase in γ to a point, then at infinity they again converge. The optimal allocation places more in the risky occupation when they are gross complements and more in the safe when they are substitutes.

$\{c_1^*(j), c_n^*(j), c_d^*(j), a^*(j), n^*(j), \ell^*(j)\}$ *satisfying Definition 3.1 for occupation-independent benefit rate b funded by occupation-independent tax rate* τ. *Then,* $\exists \tau^* > 0$ *such that*

$$\frac{\partial \int \ell(j) EU(j, \tau^*) dj}{\partial \tau} = 0 \tag{10}$$

and

$$E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))](\tau^*) > E_j[U(c_1^*(j), c_n^*(j), c_d^*(j))](0) \quad \forall j \tag{11}$$

Proof. See Online 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. Proposition 3.5 shows the optimal SDI system is unique and between zero and this extreme.

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 3.6 (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 Online Appendix B.7. □

As a Corollary to Proposition 3.6, we show the first best, planner's allocation can be achieved if occupation specific benefits are available.

Corollary 3.7 (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 4 given arbitrary occupation-specific benefits* $\hat{b} = \{b_j\}_j$ *funded by occupation-specific taxes* $\widehat{\tau} = \{\tau_j\}_j$. *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,* $\exists \mathbf{b}, \boldsymbol{\tau}$ *such that* $\mathcal{A}^*(\widehat{b}, \widehat{\tau}) = \mathcal{A}^{cm}$

Proof. See Online Appendix B.7. □

Fig. 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 Fig. 5.

The first panel in Fig. 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 after reallocation. Similarly, workers with high disability risk are actually worse off if they do not become disabled. This is because SDI causes more workers to enter the high risk occupations and lowers wages for workers in these occupations, but ex ante they gain from risk sharing. The final panel shows the total expected benefit for workers in the second period of life. The welfare benefits are non-monotone across the weighted average: they can increase as we increase the risk level and increase the number of disabled, but also decrease because the composition is not changing as much as the dynamics within disability status.

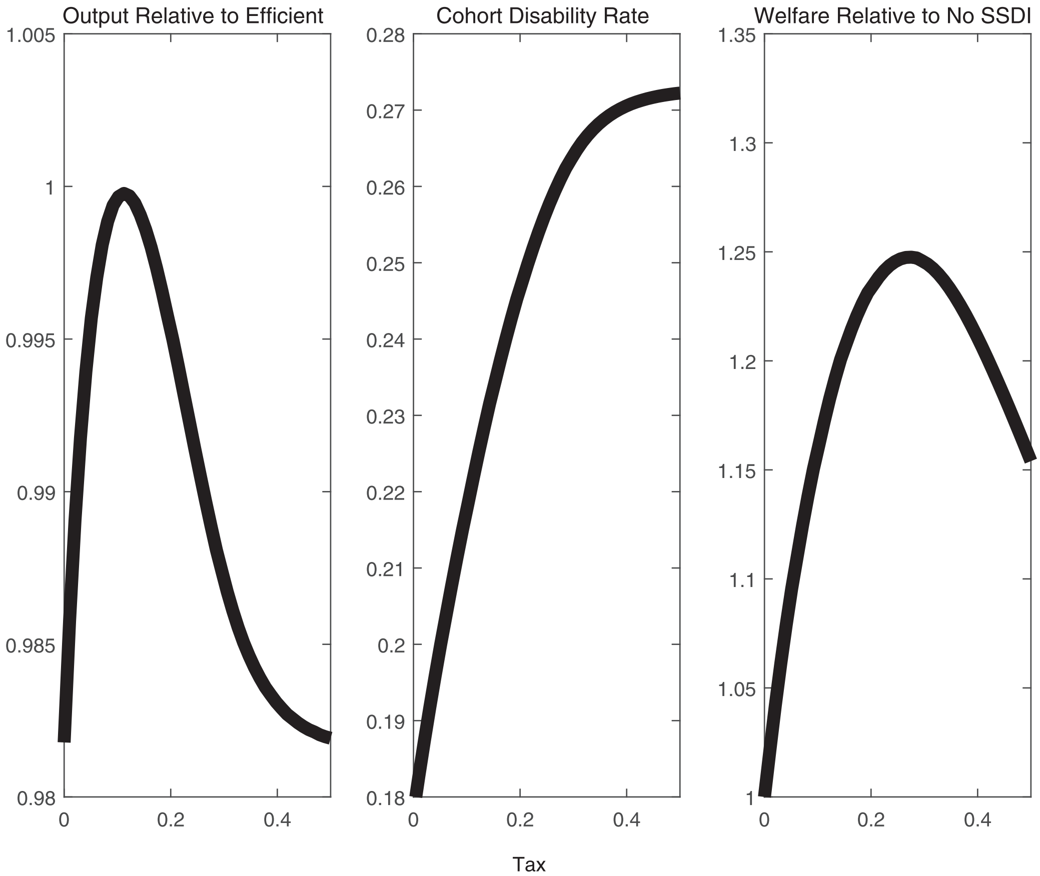

Fig. 4. Output, disability, and welfare for different tax levels. The peak output is achieved at a lower tax than the welfare maximizing tax.

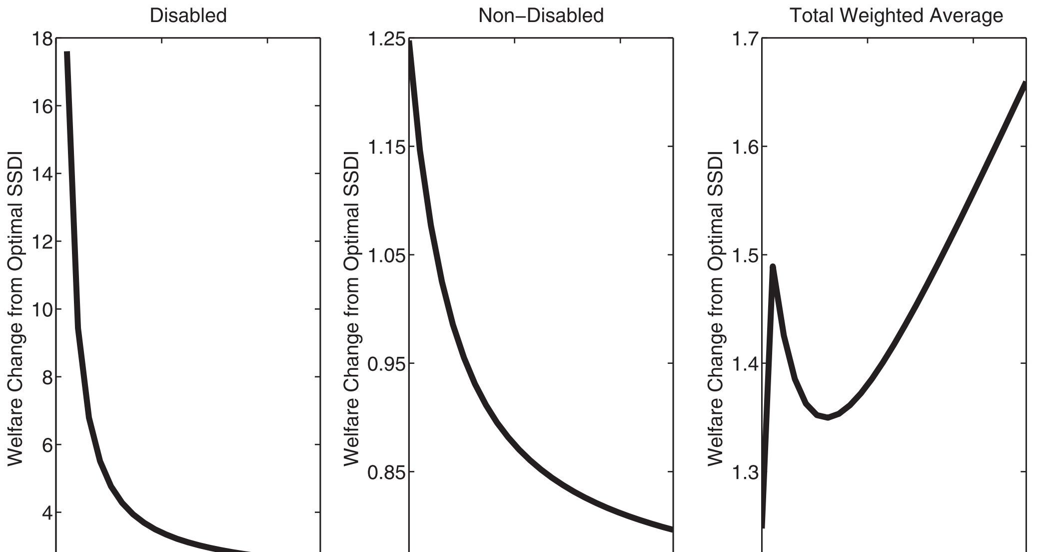

Fig. 5. The welfare effect of SDI on different groups. Among low-risk occupations, the benefit comes from increased output and therefore wages, which is seen in ex post gains. High-risk occupations gain in an ex ante sense from risk sharing.

4. 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 disability risk. We 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.

4.1. 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 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 the Online Appendix Section B.8.

Proposition 4.1 shows our assumptions on heterogeneity provide monotone occupational sorting in the form of a cut-off rule.[4] Patient, high-beta 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.

Proposition 4.1 (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. See Online Appendix B.8. □

Proposition 42 shows 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: the allocation of workers to occupations is chosen to achieve productive efficiency. This efficient distribution of workers first-order stochastically dominates the distribution in the competitive allocation. A marginal introduction of SDI improves productive efficiency. It moves the labor allocation towards more risky occupations, both within beta-types and by moving the threshold $\bar{j}(\tau)$ of occupations held by high-beta types upwards.

Proposition 4.2 (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 $\sum_j (2-\theta_j)\ell_j \tau w_j = \sum_j \theta_j \ell_j b w_j$). Then:*

$$\frac{dEU^i(\tau)}{d\tau}||_{\tau=0} > 0 \quad \forall i \in \{\ell, h\} \tag{12}$$

Proof. Online Appendix B.8. □

While all workers' welfare improves, an interesting result 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. These low-beta types also gain the least from the second channel, occupational reallocation. As the threshold $\bar{j}(\tau)$ increases the patient workers move into the riskier jobs previously held by low-beta types thus lowering their wage premium. Conversely, workers with low disability risk value the insurance benefits of SDI more than individuals with high risk and see an increase in their relative wage. This is a qualitative distinction from models of SDI that treat disability risk as common across individuals.

4.2. 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 Online Appendix Section B.9. Let preferences over consumption in the disabled state be represented by the utility function u^d. We assume:

$$(i) \quad u_d(c(1+\chi)) = u(c); \quad (ii) \quad u_d'(c(1+\chi)) \propto u'(c); \quad \chi \geq 0, \quad \forall c > 0. \tag{13}$$

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 and Pistaferri (2015). Costs of disability have constant marginal

[4] 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.

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:

$$\max_{\ell(j),c_d(j),c_n(j),c_1(j)} \int_j [u(c_1(j)) + \theta(j)u^d(c_d(j)) + (1-\theta(j))u(c_n(j))]\ell(j)dj \tag{14}$$

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

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

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:

$$\frac{\ell(k)}{\ell(j)} = \left(\frac{(2-\theta(k))}{(2-\theta(j)}\right)^{\frac{\gamma}{1-\gamma}} \left(\frac{(2+\chi\theta(k))}{(2+\chi\theta(j))}\right)^{\frac{1}{1-\gamma}} \tag{15}$$

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 a 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.

Proposition 4.3 shows our main result holds in this environment: the introduction of SDI improves welfare through both the risk sharing channel and channel of labor reallocation that increases aggregate productivity. However, this result does 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.

Proposition 4.3 (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 ($\tau \ge 0$, $b_j = bw_j$, and $\sum_j (2-\theta_j)\ell_j \tau w_j = \sum_j \theta_j \ell_j b w_j$). Then:*

$$\left.\frac{dEU(\tau)}{d\tau}\right\|_{\tau=0} > 0 \tag{16}$$

Proof. Online Appendix B.9. □

4.3. Non-verifiable disability risk and status

Here, we consider optimal SDI when neither ex-ante disability risk nor realized disability status is observable. In our baseline model, we have already restricted SDI not to condition on occupation, therefore the a scheme is robust to non-verifiable ex ante disability risk, $\theta(j)$. For the latter, truth-telling about current disability status holds below a bound on insurance. In other words, 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. Proposition 4.4 makes explicit this result in our two period economy. However, it is a concern for quantitative analysis if a large population of young workers relative to older workers at risk of disability allows full insurance.

Proposition 4.4 (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, τ), is greater than any time-consistent deviation:*

$$\begin{aligned} EU(j,\tau) &= \max_a u((1-\tau)w(j,\tau) - a) + \theta(j)u(bw(j,\tau)+a) + (1-\theta(j))u((1-\tau)w(j,\tau)+a) \\ &\ge \max_a u((1-\tau)w(j,\tau) - a) + \theta(j)u(bw(j,\tau)+a) + (1-\theta(j))\max\{u((1-\tau)w(j,\tau)+a), u(bw(j,\tau)+a)\} \end{aligned} \tag{17}$$

Proof. Online Appendix B.10. □

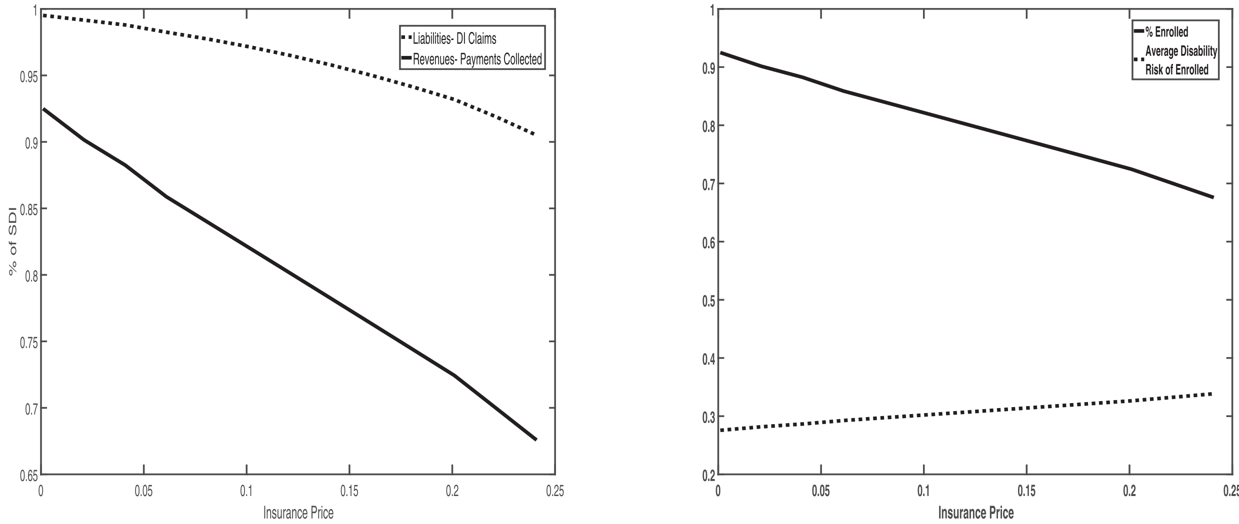

Fig. 6. As the price of disability insurance increases, the contracts remain unprofitable (Left) and the risk pool deteriorates (Right). On left, Revenue and outlays of contracts are normalized to their fraction of SDI with the same price (tax).

4.4. Private contracts

In the U.S., private disability insurance often has relatively small payouts and is not universally offered. Because of this, Golosov and Tsyvinski (2006) do not consider private contracts. Here, 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 averse selection.

The crucial constraint for private insurance is that, like social insurance, it cannot condition on the worker's occupational risk level, $\theta(j)$.[5] 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 there is over-subscription into these contracts from the 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:

$$\max_{g,a} u(w^{CM}(j) - pw^{CM}(j)g - a) + \theta(j)u(bw^{CM}(j)g + a) + (1 - \theta(j))u(w^{CM}(j) + a) \tag{18}$$

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

$$\int_j \theta(j)pw^{CM}(j) - (2 - \theta(j))bw^{CM}(j)dj = 0 \tag{19}$$

At any price, the private insurance contracts incur negative profits, as shown in the left pane of Fig. 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. Fig. 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 averse selection: the workers who chose to participate most are the worst risks.

[5] 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.). For comparison's sake, we keep the same information restrictions on both social and private insurance.

Table 3
The optimal system compared to observed U.S. system, constrained optimal policy, and actuarially fair in a calibrated model.

	Elasticity of substitution		
	$\frac{1}{5}$ (%)	2 (%)	5 (%)
Young labor in risky occupation			
Baseline	17	17	17
Constrained optimal	17.17	18.75	21.59
Actuarially fair	16.90	16.02	14.63
Welfare gain from constrained optimal DI			
(consumption equiv rel. to baseline)	+ 1.04	+ 1.04	+ 1.03
(consumption equiv rel. to actuarially fair)	+ 6.3	+ 6.3	+ 6.2
Constrained optimal output			
(rel. to baseline)	+ 0.00	+ 0.00	+ 0.00
(rel. to actuarially fair)	+ 0.00	+ 0.06	+ 0.12
Aggregate disability			
(rel. to baseline)	+ 0.14	+ 1.50	+ 3.92
(rel. to actuarially fair)	+ 0.23	+ 2.34	+ 5.95
Replacement rate			
Baseline	40.00	40.00	40.00
Constrained optimal	121.13	120.68	118.95
Tax rate			
Baseline	1.59	1.59	1.59
Constrained optimal	4.80	4.85	4.90

5. Quantitative evaluation

We now provide a simple quantitative model to explore the relative magnitudes of reallocation and risk-sharing within the policy region of the U.S. and explore policy counterfactuals.

5.1. Calibration

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).[6] We consider only two occupations, guided by the natural break in occupational risks seen in Fig. 1. We target statistics we documented associated with this break: the disability risk in the high-risk occupation is twice that of the low-risk occupation ($\pi^h = 2\pi^\ell$) and 17% of the young choose the high-risk occupation. We choose the low-risk occupation's disability hazard to be $\pi^\ell = 0.13$. This provides a disability beneficiary share of the population equal to the (demographic adjusted) average from 1985 to 2016: 3.9%.[7]

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 Autor and Duggan, 2006) and then solve for the implied balanced-budget tax.

Preferences follow (Low and Pistaferri, 2015): $u(c) = \frac{(c)^{1-\sigma}-1}{1-\sigma}$ for the non-disabled and $u_d(c) = \frac{(c*\exp(\eta))^{1-\sigma}-1}{1-\sigma}$ for the disabled. We use their calibrated parameters: $\sigma = 1.5$ following the literature, and $\eta = -0.448$ following their estimation using micro-data on consumption. 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 $\frac{1}{5}$ to 5.[8] For each γ, we recalibrate α to match 17% of workers choosing the high-risk occupation (to match the statistic we documented in the HRS data) and set Q to normalize output to 1.

5.2. Results

Using the calibrated production, preference, and occupational risk parameters, we calculate the equilibrium under 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 two counterfactual policy regimes as well as the baseline calibrated to the U.S. system. Relative to the actuarially fair

[6] This break is motivated by the fact that 71% of disabled beneficiaries were 55 years of age or older in 2016 (Authors' calculations from data released with Administration (2017)).

[7] Authors' calculation following Michaud and Wiczer (2017). The choice of $\pi^\ell = 0.13$ is not far from the analogous ADL risk of 8% in the low-risk cluster of occupations.

[8] See Katz and Autor (1999) for a literature review on the substitutability of differentiated types of labor, broadly defined. Most studies place this number between 1 and 2. We provide results for a wider range both because the dispersion of estimates is high and to better understand the quantitative properties of mechanism we are studying.

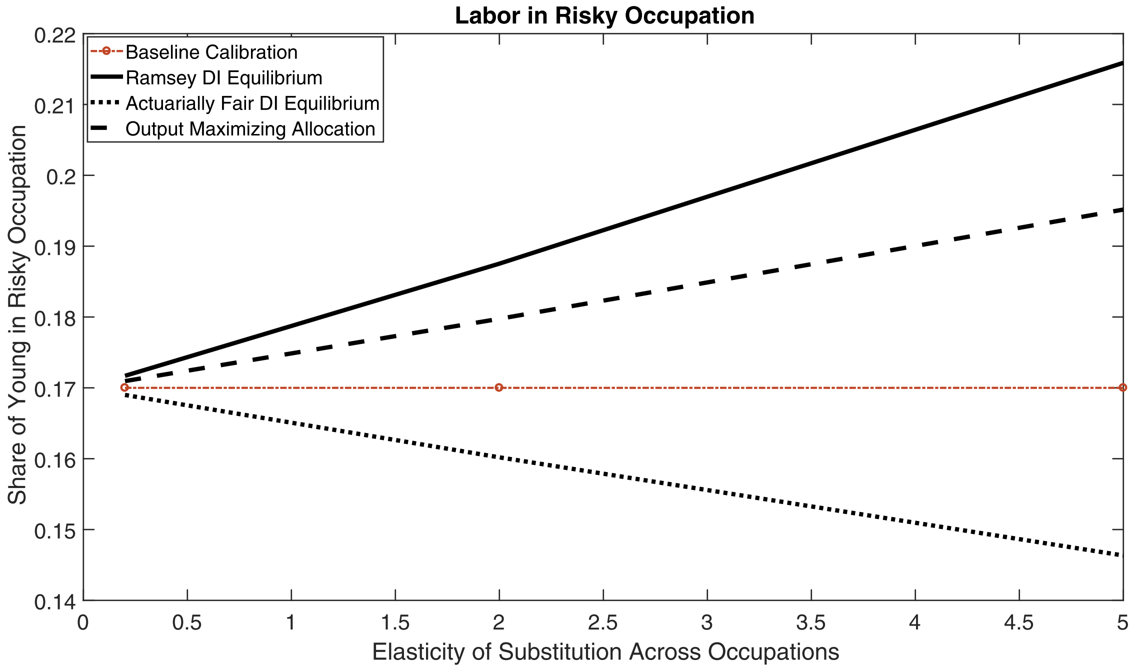

Fig. 7. Labor in the risky occupation is higher under the baseline and constrained optimal SDI programs than the output maximizing allocation. It is lower with only actuarially fair DI.

system, the constrained optimal SDI system is a substantial improvement in welfare, 6.3% in consumption equivalent terms. Further, the constrained optimal policy increases welfare over the baseline by 1.04%. This necessitates large changes in policy including a three-fold increase in the replacement rate and a doubling of the tax rate. This expanse is driven by the Low and Pistaferri (2015) calibration of the increased marginal utility of consumption for the disabled.

To help evaluate the channel of productive efficiency, Fig. 7 shows the labor allocation under each policy while changing the elasticity of substitution. The baseline target of 17% of workers in the risky occupation is fewer than the output-maximizing allocation. The constrained optimal allocation moves beyond productive efficiency, a general result proved in Proposition 3.6. However, there are negligible output effects when comparing the baseline and constrained optimal allocation. This is an artifact of the calibration of the weight α in the production technology required to match 17% of workers in the risky occupation. For very low elasticity of substitution, the calibration implies a very low output weight on high-risk occupations and therefore little effect, whereas for very high substitutability again output effects are small because workers are essentially fungible. Thus, welfare gains of the constrained optimal allocation compared to the baseline come primarily from increased risk-sharing with little output losses from moving beyond productive efficiency.

6. 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 resembles moral hazard: insuring risky occupations encourages more risk-taking, but at the margin this is efficient.

We provided 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 6.3% increase in consumption in a world with actuarially fair insurance alone and 1.04% when compared to the current policy regime. However, these gains come primarily through additional risk sharing, not productive efficiency.

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 U.S. (Isenberg, 2016). Employer-provided disability insurance is instead much more common, covering about $\frac{1}{3}$ of the workforce. This coverage, however, is rationed and occupation is a good indicator of access. 98% of workers enroll if DI is offered by their employers, but the rate at which it is offered is very heterogeneous across occupations (Groshen and Perez, 2015).[9] 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 $\theta(j)$, utilizing employers' additional information to ration insurance.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmoneco.2018.04.002.

References

Acemoglu, D., Shimer, R., 1999. Efficient unemployment insurance. J. Polit. Econ. 107 (5), 893–928.

Administration, S.S., 2017. Annual Statistical Report on the Social Security Disability Insurance Program, 2016. Technical Report, 1311826. Office of Retirement and Disability Policy: Office of Research, Evaluation and Statistics.

Autor, D.H., Duggan, M.G., 2006. The growth in the social security disability rolls: a fiscal crisis unfolding. Journal of Economic Perspectives 20 (3), 71–96. doi:10.1257/jep.20.3.71.

Fletcher, J.M., Sindelar, J.L., Yamaguchi, S., 2011. Cumulative effects of job characteristics on health. Health Econ. 20 (5), 553–570.

Golosov, M., Tsyvinski, A., 2006. Designing optimal disability insurance: a case for asset testing. J. Polit. Econ. 114 (2), 257–279.

Groshen, E.L., Perez, T.E., 2015. National Compensation Survey: Employee Benefits in the United States, March 2015. Bulletin, 2782. U.S. Bureau of Labor Statistics.

Health and Retirement Study, 2013. Core public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Health and Retirement Study, Ann Arbor, MI. http://hrsonline.isr.umich.edu/index.php?p=regcou.

Isenberg, K.N., 2016. U.S. Individual Disability Income Insurance 2015 Annual Supplement. Technical Report. LIMRA.

Katz, L.F., Autor, D., 1999. Changes in the wage structure and earnings inequality. Handb. Labor Econ. 3, 1463–1555.

Low, H., Pistaferri, L., 2015. Disability insurance and the dynamics of the incentive insurance trade-off. Am. Econ. Rev. 105 (10), 2986–3029.

Michaud, A.M., Wiczer, D., 2017. The Disability Option: Labor Market Dynamics with Macroeconomic and Health Risks. Technical Report. Mimeo.

Morefield, B., Ribar, D.C., Ruhm, C.J., 2012. Occupational status and health transitions. BE J. Econ. Anal. Policy 11 (3), 1–29.

RAND HRS Data, Version P, 2016. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. RAND HRS Santa Monica, CA. https://www.rand.org/labor/aging/dataprod/hrs-data.html.

Ravesteijn, B., Kippersluis, H.v., Doorslaer, E.v., 2018. The wear and tear on health: what is the role of occupation? Health Econ. 27 (2), e69–e86. https://www.ncbi.nlm.nih.gov/pubmed/28901590.

Schulhofer-Wohl, S., 2011. Heterogeneity and tests of risk sharing. J. Polit. Econ. 119 (5), 925–958.

[9] 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.