The Disability Option: Labor Market Dynamics with Economic and Health Risks

Amanda Michaud & David Wiczer

Data Appendix

1 Motivational Data

We motivate our analysis with figures depicting the rise in both the number and inflows of current Social Security Disability claimants. We then show this rise cannot be fully accounted for by what changes in eligibility and demographics would predict. This section describes how these figures were constructed.

Where possible, data was gathered at a disaggregated "cell" level partitioned by the cross-product of gender and seven age groups: (i) age 18-29; (ii) age 30-39; (iii) age 40-44; (iv) age 45-49; (v) age 50-54; (vi) age 55-59; and (vii) 60-64.

The following variables were collected at the cell-level spanning 1985-2014 from the SSA's *2015 Annual Statistical Supplement to the Social Security Bulletin.*

- "Total New Awards": To disabled workers only (not dependents).
- "Total Current Payees": Disabled workers only (not dependents).
- "Total Insured Workers": Estimated from the SSA's continuous work history (1%) sample. As discussed in the text, eligibility requires a certain number of "credits"- quarters meeting a minimum earnings threshold- where the number of credits required for insured status is age dependent.

Population data at the cell level was gathered from the United States *Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Bridged-Race Population Estimates,* accessed on Aug 1, 2016 2:14:55 PM.

Aggregate trends and counterfactuals in new awards and total current payees were calculated as follows:

- **Actual Trends**. Sum each series, total new awards and total current payees, across demographic cells in each year. Divide by total population age 18-64.
- **Predicted by change in eligibility alone**. Fix the new award/current payee rate as a percent of eligible for each cell at the 1985-1989 average. Predict total new award/current payees per cell by multiplying the 1985-89 rate by the actual "Total Insured Workers" divided by "Total Population" at each cell-year. Next, sum across cells in each year weighting each cell by its average total population share in 1985-89.

- **Predicted by change in eligibility and demographics**. Fix the new award/current payee rate as a percent of eligible for each cell at the 1985-1989 average. Predict total new award/current payees per cell by multiplying the 1985-89 rate by the actual "Total Insured Workers". Next, sum across cells in each year.

Figure 1 shows the results of this prediction exercise. Two facts were are interested in emerge. First, changes in demographics and eligibility account for only one-third of the rise in new awards from 1985 to present. Second, there are large fluctuations in new awards including a recent decline that cannot be accounted for by slow-moving demographics and eligibility trends.

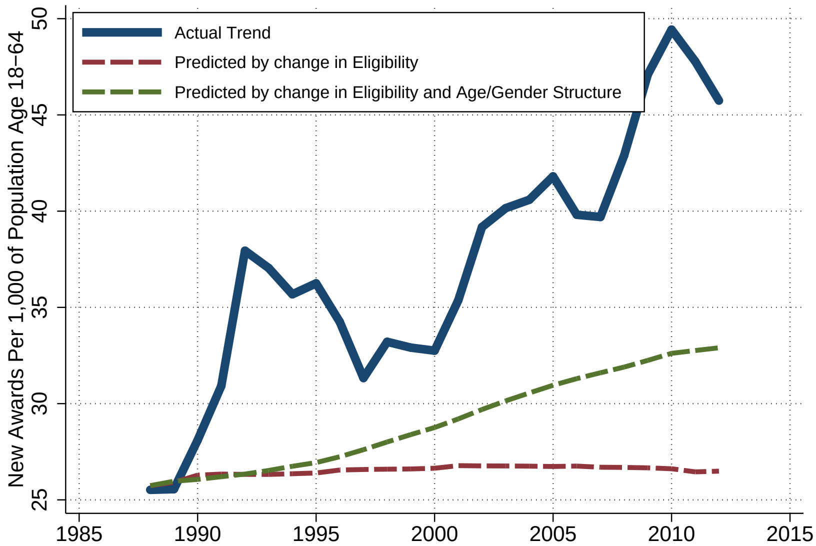


Figure 1: Predicted change in new awards (inflows to SSDI).

Another way to understand this exercise is by looking at the bar charts in Figure 2. The first row shows the change in the demographic composition of the US population and the change in the percent insured in each age-gender cell between the second half of the 1980s and the first half of the 2010s. When constructing the predicted change in new awards in Figure 1, we are fixing the award rates per insured to the 1985-89 average and shifting the demographic composition and rates insured from the black bars to the white bars; ie: from their 1985-89 levels to the 2010s. The second row of Figure 2 depicts changes in the award rate and current beneficiary status by the demographic cells. These within demographic changes provide the gap between our predicted new award rate and the actual award rate. These are the changes we are seeking to understand in this project.

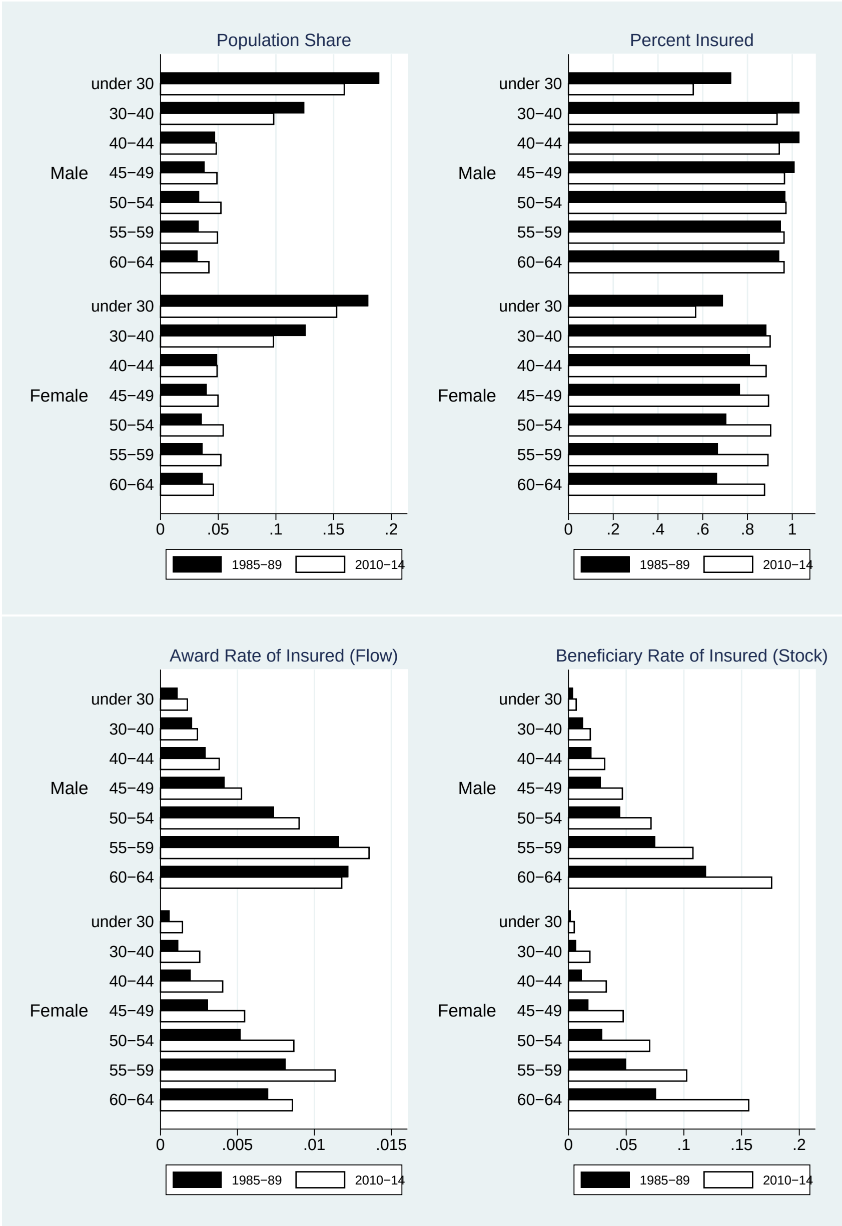


Figure 2: Changes in demographic composition and SSDI outcomes by demographic.

Although the concept of disability in our model is fairly general, our strategy to map the model to the data focuses on the physical component of disability as opposed to mental or emotional health conditions. We make this choice because we are focused on understanding trends over-time in disability awards. Figure 3 shows that the share of initial awards with a major cause of a Musculoskeletal condition have doubled since the 1980s and are now represent the largest cause of disability.

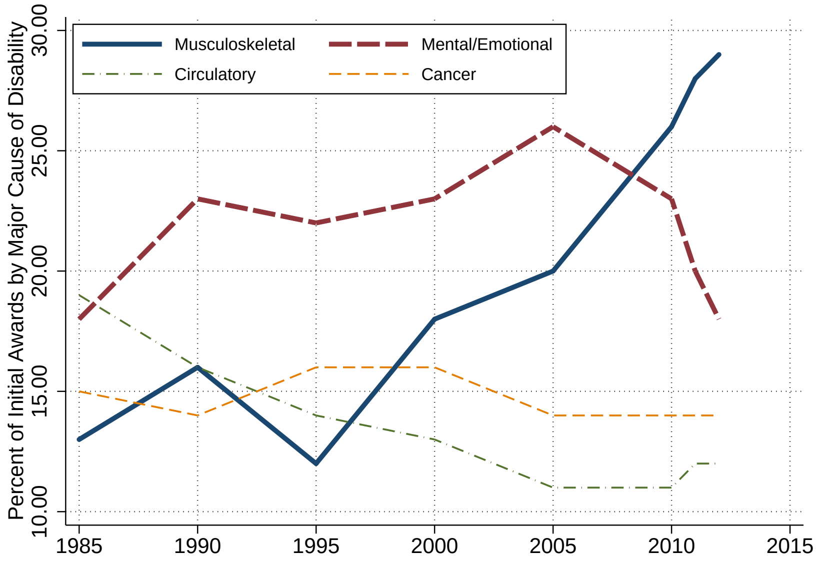


Figure 3: Changes in demographic composition and SSDI outcomes by demographic.

Figure 4 further motivates our inclusion of vocational considerations in this paper. It shows that the rise in overall new awards has occurred mostly through a rise in awards with vocational considerations.

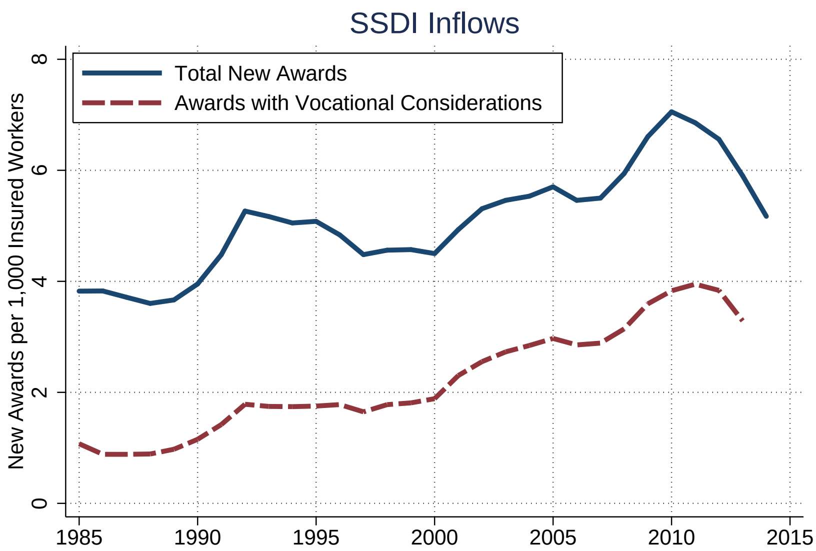


Figure 4: Awards with Vocational Considerations drive the increase in overall awards

Figure 5 motivates our focus on matching flows onto DI rather than the stock of current DI beneficiaries. It shows that the exit rate from DI has decreased substantially overtime. This is a channel increasing the stock that the model only speaks partially to. The model will predict that younger and more healthy individuals will be entering DI overtime, thus lowering the average exit rate via death or retirement shown in this graph. However, it cannot fully account for the decline in the death rate, likely through missing features such as improvements in medical technologies.

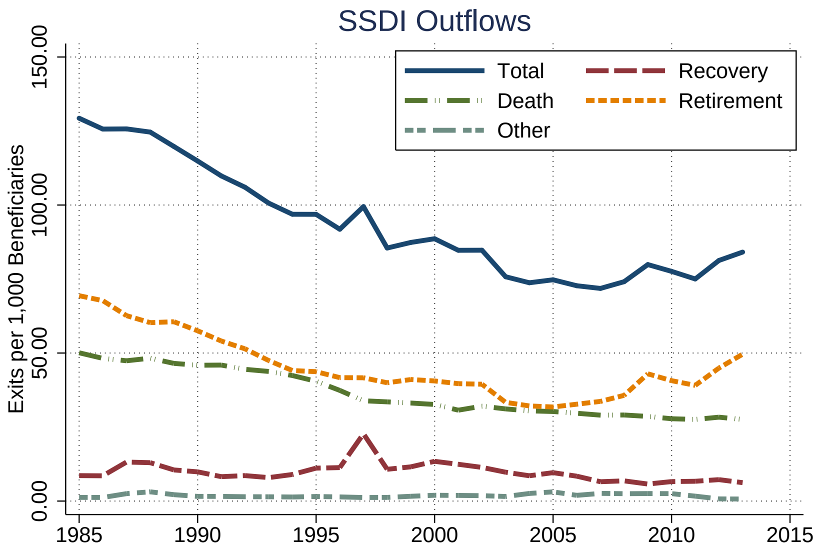


Figure 5: Exit rate of SSDI beneficiaries by reason

2 Panel Study of Income Dynamics (PSID)

We use the PSID to analyze various aspects of labor market and health dynamics and their relationships with one another. Throughout our sample is limited to males.[1] We keep the SEO sample as it is disproportionably low-income, a relevant population for this study. We drop the latino sample and respondents that we see fewer than 3 times. This section begins by explaining variable construction and then provides additional calculations of statistics used in the text for alternative sample design as a robustness check.

2.1 Sample and Variable Construction

Health Statistics. We replicate health status coding from Low et al. (2015).[2] We use three questions asked in the PSID starting with: (i) "Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?" If the respondent answers affirmatively, the next question asked is: (ii) "Does this condition keep you from doing some types of work?". To this question there are three possible responses: "Yes," "No," or "Can do nothing." Those answering either of the two former responses are then asked the third question:

[1] This is because health related questions pertaining to all members of the household are answered only by the head of household (often male). Prior studies have shown reporting on others' health introduces bias that we cannot easily correct for. However, we do provide estimates for the entire sample including females for select statistics in this appendix as they may be of interest to the audience.

[2] We refer the reader to their paper which validates this measure against alternative datasets.

(iii) "For work you can do, how much does it limit the amount of work you can do?" To this question the possible answers are: "A lot," "Somewhat," "Just a little," or "Not at all."

Health status takes three values: "No work limitation" ($d = 0$); "Moderate work limitation" ($d = 1$); and "Severe limitation" ($d = 2$). Respondents are coded as having no work limitation if they answer "No" to question one or "Not at all" to question three. They are coded as having a moderate limitation if they answer "Yes" to question one and "No" to question two OR "Yes" to question two and "Somewhat" or "Just a little" to question three. The remainder answering "yes" to question one are coded as having a severe limitation.

Labor Statistics. We use the PSID calculated hourly wage variable, available all years except 1993. We deflate this value to 1999 US dollars using the CPI-U multiplier from the Bureau of Labor Statistics. We drop top-coded values and values below $3.00. We define employment as answering either "working right now" or "only temporarily laid off". Where used, "full-time full-year" employment refers to those usually working more than 30hrs per week for at least 50 weeks per year; or at least 1500 hours per year.

Lifetime Occupation. We consider several definitions of "lifetime" occupation and also perform robustness considering "most recent" occupation instead. First, we define occupations in the 16 SOC codes. Prior to 2003, respondents provide a single occupation and we use this as their occupation for the year. From 2003 onwards, respondents report occupation and earnings for up to three jobs. In these years, we code the job with the highest earnings as the respondent's occupation for the year. From here, we compute the modal occupation in which we most often view the respondent over the entirety of the available panel. When we must break a tie, we choose the higher SOC value. This is because the lower SOC values are most associated with career progression to managerial and professional occupations; ie: the respondent may be manager over the same type of occupation. Next we make a decision about whether the respondent has been in his occupation for long enough for us to code it as a lifetime occupation. We drop individuals who do not meet this criteria. To make this judgement call we use three variables. First, the number of time the individual is observed in their modal occupation. Second, the max employer tenure reported by the individual while in that occupation. Third, the individual's answer to the question: "Have you had a number of different kinds of jobs, or have you mostly worked in the same occupation you started in, or what?". There are three possible responses: (i) "Have had a number of different kinds of jobs"; (ii) "Both; have had a number of different jobs but mostly the same occupation"; and (iii) "Mostly the same occupation". We consider respondents to have self-reported working "mostly in the same occupation" if they answer (ii) or (iii) AND are over the age of 39. We consider the following four specifications, the first of which is the most inclusive and used in the main tex.

- Lifetime Occ1: Observed more than 4 years in the same SOC; or reports job tenure greater than 4 years; or reports working "mostly in the same occupation".

  - This drops 449 individuals (1.4% of the sample) who report occupation at some point, but do not meet the criteria for having a lifetime occupation.[3]
  - 77.5% of the sample over (1983-1998) have a current occupation that matches their lifetime occupation; 76.4% for the full 1983-2013.
  - 77.7% of those over 50 and younger than 63 over (1983-1998) have a current occupation that matches their lifetime occupation; 74.1% for the full 1983-2013.

- Lifetime Occ2: Observed more than 9 years in the same SOC; or reports job tenure greater than 9 years.
  - This drops 4,807 individuals (14.6% of the sample) who report occupation at some point, but do not meet the criteria for having a lifetime occupation.
  - 81.9% of the sample over (1983-1998) have a current occupation that matches their lifetime occupation; 80.2% for the full 1983-2013.
  - 80.2% of those over 50 and younger than 63 over (1983-1998) have a current occupation that matches their lifetime occupation; 80.3% for the full 1983-2013.
- Lifetime Occ3: All those who report working "mostly in the same occupation".
  - This drops 17,894 individuals (54.5% of the sample) who report occupation at some point, but do not meet the criteria for having a lifetime occupation.
  - 83.35% of the sample over (1983-1998) have a current occupation that matches their lifetime occupation; 85.8% for the full 1983-2013.
  - 83.9% of those over 50 and younger than 63 over (1983-1998) have a current occupation that matches their lifetime occupation; 83.0% for the full 1983-2013.
- Lifetime Occ4: Current or most recent occupation.

Where necessary, we use the following bridge to harmonize data with occupations coded in 1990 Census codes to SOC codes used in the HRS.

- $SOC = 1$: "Managerial specialty" (Census 1990: 003-037)
- $SOC = 2$: "Professional specialty operation and technical support" (Census 1990: 043-235)
- $SOC = 3$: "Sales" (Census 1990: 243-285)
- $SOC = 4$: "Clerical, administrative support" (Census 1990: 303-389)
- $SOC = 5$: "Service: private household, cleaning and building services" (Census 1990: 403-407)
- $SOC = 6$: "Service: protection" (Census 1990: 413-427)
- $SOC = 7$: "Service: food preparation" (Census 1990: 433-444)

[3]Recall, we already drop all individuals seen fewer than three times.

- $SOC = 8$: "Health services" (Census 1990: 445-447)
- $SOC = 9$: "Personal services" (Census 1990: 448-469)
- $SOC = 10$: "Farming, forestry, fishing" (Census 1990: 473-499)
- $SOC = 11$: "Mechanics and repair" (Census 1990: 503-549)
- $SOC = 12$: "Construction trade and extractors" (Census 1990: 553-617)
- $SOC = 13$: "Precision production" (Census 1990: 633-699)
- $SOC = 14$: "Operators: machine" (Census 1990: 703-799)
- $SOC = 15$: "Operators: transport, etc." (Census 1990: 803-859)
- $SOC = 16$: "Operators: handlers, etc." (Census 1990: 863-889)

Other Variables. We code three education groups corresponding to years of schooling: less than high school = 11 years or less; high school = 12 years; college = more than 12 years. Our 5 age categories correspond to the model: Age=1 are 30-45; Age=2 are 46-55; Age=3 are 56-60; Age=4 are 61-63; and where applicable "old" is over 65 (used for death probability only).

SSDI Enrollment. The PSID has a question directly asking if the head receives SSDI income in only years 1986-1993, 2005, 2007, and 2011. In the available years, we use this question directly, coding gaps between two DI years as a DI year. For years in between we impute that the head receives SSDI if they or their family reports Social Security income AND the answer that they are not at work due to a disability. We provide statistics for both the imputed variable and the years with the direct question, separately.

2.2 Summary Statistics.

The following tables provide prior labor market statistics for three groups of individuals: those receiving DI; those in the reference population (in the sample, aged 45-60, non-college); and by work limitation status. We construct an indicator that equals one for each of the following if they are ever observed in any of the four years prior to the current survey year: labor income less than 20th percentile of non-college in that year; involuntary separation; and involuntary separation currently unemployed. The following tables report the share for which this indicator variable is positive.[4]

[4] For labor income, we only include those actually earning positive labor income.

Table 1: Labor income less than 20th percentile in any of past 5 years; DI Beneficiaries and reference pop.

	1986-2013 DI Beneficiaries	1986-1993 DI Beneficiaries	2005, 2007, 2011 DI Beneficiaries	1986-1993 Reference Pop	2005, 2007, 2011 Reference Pop
No	0.2279 **	0.0892 *	0.3078 **	0.8007 **	0.7799 **
	(0.0329)	(0.0349)	(0.0427)	(0.0077)	(0.0104)
Yes	0.7721 **	0.9108 **	0.6922 **	0.1993 **	0.2201 **
	(0.0329)	(0.0349)	(0.0427)	(0.0077)	(0.0104)
Observations	435	153	178	6046	2488

Col 1 includes imputation
Standard errors in parentheses
† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 2: Labor income less than 20th percentile in any of past 5 years; by work limitation.

	1986-1993 moderate	2005, 2007, 2011 moderate	1986-1993 severe	2005, 2007, 2011 severe
No	0.4037 **	0.2035 **	0.2575 **	0.1009 **
	(0.0229)	(0.0295)	(0.0258)	(0.0273)
Yes	0.5963 **	0.7965 **	0.7425 **	0.8991 **
	(0.0229)	(0.0295)	(0.0258)	(0.0273)
Observations	887	408	392	140

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

Table 3: Involuntary Separation in any of past 5 years. DI Beneficiaries and Reference Pop

	1986-2013 DI Beneficiaries	1986-1993 DI Beneficiaries	2005, 2007, 2011 DI Beneficiaries	1986-1993 Reference Pop	2005, 2007, 2011 Reference Pop
No	0.9427 **	0.8623 **	0.9309 **	0.8189 **	0.8710 **
	(0.0102)	(0.0246)	(0.0152)	(0.0072)	(0.0073)
Yes	0.0573 **	0.1377 **	0.0691 **	0.1811 **	0.1290 **
	(0.0102)	(0.0246)	(0.0152)	(0.0072)	(0.0073)
Observations	1562	538	463	7624	3413

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

Table 4: Involuntary Separation in any of past 5 years, by work limitation

	1986-1993 moderate	2005, 2007, 2011 moderate	1986-1993 severe	2005, 2007, 2011 severe
No	0.8007 **	0.8418 **	0.8653 **	0.8800 **
	(0.0179)	(0.0195)	(0.0268)	(0.0267)
Yes	0.1993 **	0.1582 **	0.1347 **	0.1200 **
	(0.0179)	(0.0195)	(0.0268)	(0.0267)
Observations	1091	852	280	258

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

Table 5: Involuntary Unemployed in any of past 5 years. DI Beneficiaries and Reference Pop

	1986-2013 DI Beneficiaries	1986-1993 DI Beneficiaries	2005, 2007, 2011 DI Beneficiaries	1986-1993 Reference Pop	2005, 2007, 2011 Reference Pop
No	0.9545 **	0.8962 **	0.9442 **	0.9268 **	0.9462 **
	(0.0094)	(0.0220)	(0.0172)	(0.0047)	(0.0050)
Yes	0.0455 **	0.1038 **	0.0558 **	0.0732 **	0.0538 **
	(0.0094)	(0.0220)	(0.0172)	(0.0047)	(0.0050)
Observations	1562	538	346	7624	3413

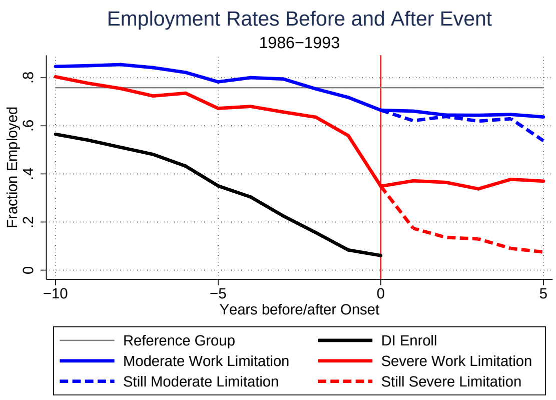
Standard errors in parentheses
† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

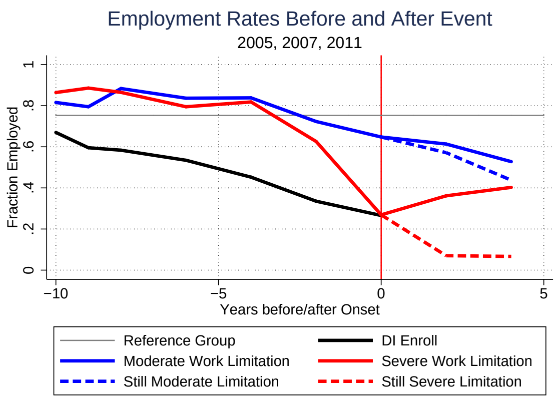
Table 6: Involuntary Unemployed in any of past 5 years. by work limitation

	1986-1993 moderate	2005, 2007, 2011 moderate	1986-1993 severe	2005, 2007, 2011 severe
No	0.9185 **	0.8856 **	0.9594 **	0.9194 **
	(0.0122)	(0.0172)	(0.0128)	(0.0228)
Yes	0.0815 **	0.1144 **	0.0406 **	0.0806 **
	(0.0122)	(0.0172)	(0.0128)	(0.0228)
Observations	1091	852	280	258

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

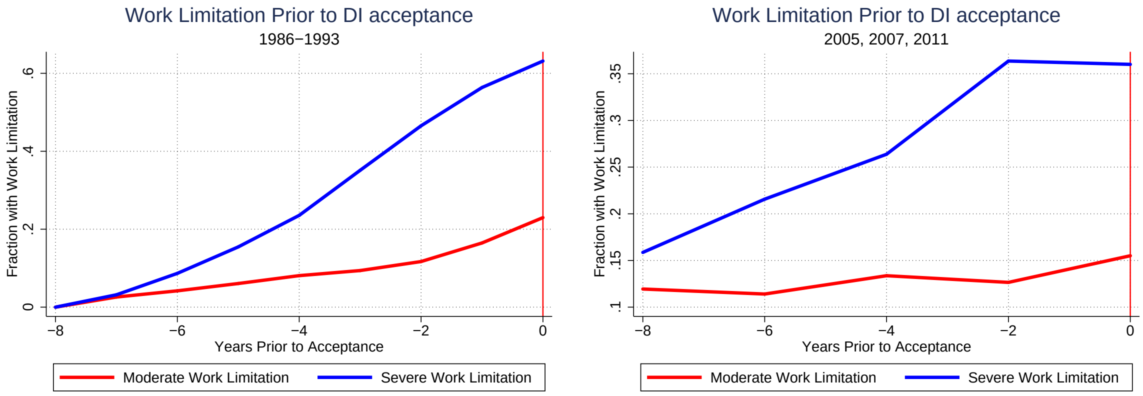
The following two graphs show employment rates in the years preceding and following the onset of a work limitation of each degree and the first year of DI receipt. The reference population include those otherwise satisfying sample criteria who are additionally age 40-62 and do not have education beyond high school.





The following two graphs show work limitation prevalence in the years preceding the first year of DI receipt. The reference population include those otherwise satisfying sample criteria who are additionally age 40-62 and do not have education beyond high school.

Another way we relate the model predictions to the data is by running a logistic regression to calculate how individuals' flows onto SSDI in the following year relate to economic conditions this year. These economic conditions contain our trend measure: wage decline as predicted by



time trends and payment to occupation tasks; and the aggregate unemployment rate.[5] Additional controls include the other key features of our model: dummies for each age and work limitation group. The regression results are presented in Table 7.

Table 7: Relationship between Economic Shocks and Flows onto DI (PSID)

Logistic Regression

	New DI $(t+1)$
Predicted Occupation Wage Trend	-1.0420 *
	(0.5116)
log(Unemployment Rate)	5.8930 **
	(1.3093)
Age 40-54	0.6272 **
	(0.1262)
Age 55-59	1.3045 **
	(0.1320)
Age 60-63	1.6975 **
	(0.1695)
Moderate Work Limitation	2.1579 **
	(0.1417)
Severe Work Limitation	4.1576 **
	(0.1152)
Constant	3.9322 *
	(2.0021)
Observations	17760

Standard errors in parentheses
† $p < 0.10$, *
Occupation wage trend predicted by O*net tasks, cubic in time, and their interactions.

2.3 Comparison to Health and Retirement Survey (HRS)

We compare our disability outcomes by longest held occupation to the Health and Retirement Survey. The health and retirement survey has the advantage of having a larger cross-section of older workers and hence a larger number of observations of individuals receiving SSDI. It also asks a question about individuals' longest held occupations. The disadvantage is that it is

[5]Since our model only contains males aged 30-63, we also use this group to calculate unemployment.

not a long panel. Therefore, we cannot see as rich labor market history data prior to disability onset. We provide one measure of health decline: the share in the occupation that report any difficulty in "Activities of Daily Living" (ADLs). We also report the share that we observe receiving SSDI. All statistics are reported for individuals we see after age 50.

		HRS		PSID 1986-1993	
SOC	NAME	ADL risk	DI risk	Severe Risk	DI Risk
1	Managerial	0.069	0.075	0.130	0.024
2	Professional	0.062	0.076	0.137	0.042
3	Sales	0.068	0.095	0.186	0.014
4	Clerical, admin	0.072	0.087	0.104	0.030
5	Service: clean/maint	0.150	0.103	n.s	n.s
6	Service: protect	0.091	0.093	0.345	0.034
7	Service: food	0.110	0.184	0.294	0.118
8	Service: health	0.130	0.144	n.s	n.s
9	Service: personal	0.142	0.169	0.205	0.124
10	Farm, fish, forest	0.080	0.145	0.161	0.071
11	Mechanics	0.129	0.185	0.231	0.096
12	Construction/extractors	0.111	0.181	0.235	0.071
13	Precision production	0.101	0.130	0.170	0.038
14	Operators: machine	0.128	0.200	0.245	0.122
15	Operators: transport	0.134	0.200	0.230	0.137
16	Operators: handlers	0.117	0.205	0.274	0.155

3 Calibration Targets

3.1 Health Transition Matrix

We calibrate the health transition matrix using the constructed work limitation variable in our PSID sample. For this section we limit our analysis to the annual data available prior to the conversion of the PSID to biannual after 1998.

The raw distribution of work limitations by age is as follows:

Table 8: Health Distribution by Age

Age Group	None	Moderate	Severe
30-45	0.913974	0.055419	0.030607
46-55	0.85285	0.088063	0.059086
56-60	0.8022	0.114454	0.083345
61-65	0.748868	0.14222	0.108912

This distribution is generated by an individual specific transition matrix in our model. Common to all individuals is a baseline risk of worsening health that is dependent on age. At the beginning of life individuals choose an occupation and draw an additional health risk (may be negative) from an occupation-specific distribution. This is added to the common age-dependent risk to calculate the individual's total risk in each stage of life. The mean of the occupation specific distribution is chosen to match a linear probability model. For a given state, we consider each transition unilaterally.[6] However, we must be careful to control for selection on individual specific factors. To do so, we use the IV strategy developed in Michaud and Wiczer (2014). Namely, we summarize the health risk component of an occupation by the intensity of physical tasks in that occupation. We IV for selection into the occupation using other non-physical tasks (See Michaud and Wiczer (2014) for detail). We also include four age group dummies corresponding to the age groups held constant through the calibration. The resulting estimates for each transition are:

Table 9: Health Transition Hazard (Linear Probability): Lifetime Occupation Spec. 1

	0-1	0-2	0-d	1-0	1-2	1-d	2-0	2-1	2-d
OccPhys	0.0031 **	0.0015 **		0.0247 †	0.0162 †		0.0044	-0.0282 †	
	(0.0007)	(0.0004)		(0.0142)	(0.0098)		(0.0118)	(0.0169)	
ageD2	0.0049 *	0.0013	0.0019 **	-0.0981 **	0.0300	0.0012	-0.1135 **	-0.0960 *	0.0027
	(0.0019)	(0.0010)	(0.0007)	(0.0371)	(0.0239)	(0.0050)	(0.0412)	(0.0484)	(0.0102)
ageD3	0.0095 **	0.0023	0.0093 **	-0.0586	0.0585 †	0.0118	-0.1417 **	-0.1057 *	0.0136
	(0.0031)	(0.0016)	(0.0020)	(0.0483)	(0.0342)	(0.0107)	(0.0383)	(0.0484)	(0.0118)
ageD4	0.0234 **	0.0086 **	0.0087 **	-0.1144 **	0.1696 **	0.0038	-0.1358 **	-0.1075 *	0.0321 †
	(0.0043)	(0.0026)	(0.0021)	(0.0408)	(0.0364)	(0.0067)	(0.0384)	(0.0491)	(0.0176)
old	0.0000	0.0000	0.0274 **	0.0000	0.0000	0.0464 **	0.0000	0.0000	0.1003 **
	(.)	(.)	(0.0026)	(.)	(.)	(0.0097)	(.)	(.)	(0.0139)
Constant	0.0123 **	0.0039 **	0.0009 **	0.3940 **	0.0912 **	0.0038	0.2182 **	0.3096 **	0.0076
	(0.0008)	(0.0005)	(0.0002)	(0.0221)	(0.0126)	(0.0027)	(0.0312)	(0.0356)	(0.0055)
Observations	42027	42027	49586	1352	1352	2261	850	850	1950

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

These regression results imply a baseline risk (for the young age group 30-45) for each occupation. This risk is calculated for each type of transition and point estimates are listed in the following table.

[6]For example, the probability of moving from moderate disability to death or moving to severe disability is independent from the probability of moving back to no disability. Therefore, we do not choose a competing hazards model because we do not consider death to censor the probability of recovery.

Table 10: Occupation Component of Health Matrix

SOC	base_0_1	base_0_2	base_1_0	base_1_2	base_2_0	base_2_1
1	0.007582	0.001605	0.356219	0.066335	0.211472	0.352771
2	0.010217	0.0029	0.377368	0.080269	0.215254	0.328572
3	0.009068	0.002336	0.36815	0.074196	0.213606	0.33912
4	0.0096	0.002597	0.372421	0.07701	0.214369	0.334233
5	0.013307	0.00442	0.40218	0.096616	0.219691	0.300183
6	0.014793	0.00515	0.414115	0.104479	0.221825	0.286527
7	0.015325	0.005412	0.418387	0.107294	0.222589	0.281639
8	0.012192	0.003872	0.393231	0.09072	0.218091	0.310423
9	0.012679	0.004111	0.397138	0.093294	0.21879	0.305952
10	0.015478	0.005487	0.419611	0.1081	0.222808	0.280238
11	0.016091	0.005788	0.424534	0.111343	0.223689	0.274606
12	0.016825	0.006149	0.430429	0.115227	0.224743	0.26786
13	0.014345	0.00493	0.410519	0.10211	0.221182	0.290641
14	0.015621	0.005557	0.420758	0.108856	0.223013	0.278926
15	0.014087	0.004803	0.408447	0.100745	0.220812	0.293012
16	0.017169	0.006318	0.433187	0.117045	0.225236	0.264704

The following tables present regression outcomes of the linear probability model for our alternative life-time occupation specifications (1-3) and the most recent occupation specification (4).

Table 11: Health Transition Hazard (Linear Probability): Lifetime Occupation Spec. 2

	0-1	0-2	0-d	1-0	1-2	1-d	2-0	2-1	2-d
OccPhys	0.0030 **	0.0012 **		0.0301 *	0.0093		0.0118	-0.0374 †	
	(0.0007)	(0.0004)		(0.0152)	(0.0101)		(0.0141)	(0.0194)	
ageD2	0.0042 *	0.0020 †	0.0019 **	-0.0937 *	0.0226	0.0012	-0.1500 **	-0.2128 **	0.0027
	(0.0019)	(0.0011)	(0.0007)	(0.0393)	(0.0233)	(0.0050)	(0.0562)	(0.0612)	(0.0102)
ageD3	0.0086 **	0.0030 †	0.0093 **	-0.0614	0.0570 †	0.0118	-0.1904 **	-0.1672 **	0.0136
	(0.0032)	(0.0017)	(0.0020)	(0.0512)	(0.0340)	(0.0107)	(0.0522)	(0.0632)	(0.0118)
ageD4	0.0203 **	0.0095 **	0.0087 **	-0.1122 *	0.1831 **	0.0038	-0.1760 **	-0.2019 **	0.0321 †
	(0.0042)	(0.0027)	(0.0021)	(0.0442)	(0.0389)	(0.0067)	(0.0527)	(0.0628)	(0.0176)
old	0.0000	0.0000	0.0274 **	0.0000	0.0000	0.0464 **	0.0000	0.0000	0.1003 **
	(.)	(.)	(0.0026)	(.)	(.)	(0.0097)	(.)	(.)	(0.0139)
Constant	0.0119 **	0.0032 **	0.0009 **	0.3887 **	0.0772 **	0.0038	0.2677 **	0.3981 **	0.0076
	(0.0008)	(0.0005)	(0.0002)	(0.0241)	(0.0125)	(0.0027)	(0.0453)	(0.0503)	(0.0055)
Observations	35975	35975	49586	1106	1106	2261	583	583	1950

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

Table 12: Health Transition Hazard (Linear Probability): Lifetime Occupation Spec. 3

	0-1	0-2	0-d	1-0	1-2	1-d	2-0	2-1	2-d
OccPhys	0.0030 **	0.0015 **		0.0236	0.0001		0.0133	-0.0330	
	(0.0009)	(0.0005)		(0.0216)	(0.0159)		(0.0140)	(0.0244)	
ageD2	0.0036	0.0017	0.0019 **	-0.0858	0.0245	0.0012	-0.0682	-0.1864 *	0.0027
	(0.0024)	(0.0012)	(0.0007)	(0.0537)	(0.0354)	(0.0050)	(0.0544)	(0.0764)	(0.0102)
ageD3	0.0042	0.0023	0.0093 **	0.0127	0.0298	0.0118	-0.0529	-0.1983 *	0.0136
	(0.0034)	(0.0020)	(0.0020)	(0.0715)	(0.0495)	(0.0107)	(0.0566)	(0.0776)	(0.0118)
ageD4	0.0197 **	0.0116 **	0.0087 **	-0.0621	0.1282 **	0.0038	-0.0866 †	-0.2271 **	0.0321 †
	(0.0051)	(0.0035)	(0.0021)	(0.0606)	(0.0487)	(0.0067)	(0.0502)	(0.0771)	(0.0176)
old	0.0000	0.0000	0.0274 **	0.0000	0.0000	0.0464 **	0.0000	0.0000	0.1003 **
	(.)	(.)	(0.0026)	(.)	(.)	(0.0097)	(.)	(.)	(0.0139)
Constant	0.0113 **	0.0031 **	0.0009 **	0.3696 **	0.0999 **	0.0038	0.1469 **	0.4198 **	0.0076
	(0.0012)	(0.0006)	(0.0002)	(0.0363)	(0.0223)	(0.0027)	(0.0440)	(0.0614)	(0.0055)
Observations	19611	19611	49586	598	598	2261	418	418	1950

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

Table 13: Health Transition Hazard (Linear Probability): Most Recent Occupation

	0-1	0-2	0-d	1-0	1-2	1-d	2-0	2-1	2-d
OccPhys	0.0021 **	0.0009 **		0.0185	0.0127		0.0078	-0.0558 **	
	(0.0007)	(0.0003)		(0.0151)	(0.0103)		(0.0146)	(0.0197)	
ageD2	0.0050 **	0.0020 *	0.0019 **	-0.1020 *	0.0339	0.0012	-0.1270 **	-0.1503 **	0.0027
	(0.0019)	(0.0010)	(0.0007)	(0.0397)	(0.0266)	(0.0050)	(0.0446)	(0.0523)	(0.0102)
ageD3	0.0090 **	0.0031 †	0.0093 **	-0.0754	0.0639 †	0.0118	-0.1527 **	-0.1268 *	0.0136
	(0.0031)	(0.0016)	(0.0020)	(0.0512)	(0.0371)	(0.0107)	(0.0434)	(0.0550)	(0.0118)
ageD4	0.0225 **	0.0092 **	0.0087 **	-0.1403 **	0.1841 **	0.0038	-0.1603 **	-0.1569 **	0.0321 †
	(0.0044)	(0.0027)	(0.0021)	(0.0432)	(0.0396)	(0.0067)	(0.0416)	(0.0529)	(0.0176)
old	0.0000	0.0000	0.0274 **	0.0000	0.0000	0.0464 **	0.0000	0.0000	0.1003 **
	(.)	(.)	(0.0026)	(.)	(.)	(0.0097)	(.)	(.)	(0.0139)
Constant	0.0114 **	0.0028 **	0.0009 **	0.4114 **	0.0951 **	0.0038	0.2384 **	0.3628 **	0.0076
	(0.0008)	(0.0004)	(0.0002)	(0.0234)	(0.0136)	(0.0027)	(0.0347)	(0.0398)	(0.0055)
Observations	39901	39901	49586	1213	1213	2261	752	752	1950

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

3.2 Employment Probability Regression

The following table displays results for a probit regression in our PSID sample. The dependent variable equals one if the individual is employed and zero otherwise. The sample includes individuals aged 30-65, seen 3 times and at least one time employed between 1983 and 1996. The first three columns provide estimates using different exclusion restrictions. The first uses 5-year change in aggregate log- full-time full-year employment for the age-education-occupation demographic of the individual. The second uses just age-education and the third uses just education. All of these statistics are calculated from the Current Population Survey and the construction is detailed in the "Current Population Survey" section of this appendix. The fourth column repeats specification (1), but includes women in the sample. The "preferred" specification used in the main text is specification #2.

Table 14: Est. Coefficients of Probit Estimation- Dependent Variable=1 if Employed. Lifetime Occupation Spec. 1

	(1)	(2)	(3)	(4)
Moderate WL	-0.9141 **	-0.9122 **	-0.9077 **	-0.9451 **
	(0.0477)	(0.0477)	(0.0476)	(0.0389)
Severe WL	-2.1419 **	-2.1302 **	-2.1236 **	-2.0253 **
	(0.0603)	(0.0604)	(0.0604)	(0.0508)
ageD2	-0.1677 **	-0.1456 **	-0.1378 **	-0.0590 †
	(0.0371)	(0.0373)	(0.0374)	(0.0327)
ageD3	-0.7304 **	-0.6477 **	-0.7095 **	-0.5723 **
	(0.0505)	(0.0537)	(0.0511)	(0.0428)
ageD4	-0.9910 **	-0.9032 **	-0.9564 **	-0.8661 **
	(0.0715)	(0.0738)	(0.0725)	(0.0594)
nWhite	-0.3104 **	-0.2860 **	-0.2646 **	-0.3642 **
	(0.0352)	(0.0354)	(0.0357)	(0.0298)
married	0.4694 **	0.4686 **	0.4619 **	0.4267 **
	(0.0420)	(0.0420)	(0.0421)	(0.0419)
5 year diff of AgeXEdXOcc FTFY Empl	0.0852			0.0913 *
	(0.0569)			(0.0357)
1 year diff of AgeXEdXOcc FTFY Empl	0.0210			-0.0253
	(0.0599)			(0.0334)
5 year diff of AgeXEd FTFY Empl		0.7788 **		
		(0.1282)		
1 year diff of AgeXEd FTFY Empl		-0.4577 †		
		(0.2755)		
5 year diff of FTFY Ed Emp			1.2054 **	
			(0.1601)	
1 year diff of FTFY Ed Emp			-0.1720	
			(0.4301)	
Female				-0.1430 **
				(0.0481)
Constant	1.4338 **	1.3791 **	1.3617 **	1.4344 **
	(0.0429)	(0.0439)	(0.0440)	(0.0418)
Observations	32080	32092	32112	41844

Standard errors in parentheses

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

These estimates translate to the following marginal effects:

Table 15: Marginal Effects of Probit Estimation- Dependent Variable=1 if Employed. Lifetime Occupation Spec. 1

	(1)	(2)	(3)	(4)
Moderate WL	-0.1982 **	-0.1967 **	-0.1940 **	-0.2441 **
Severe WL	-0.6537 **	-0.6488 **	-0.6451 **	-0.6504 **
ageD2	-0.0229 **	-0.0195 **	-0.0183 **	-0.0099 †
ageD3	-0.1422 **	-0.1201 **	-0.1349 **	-0.1253 **
ageD4	-0.2322 **	-0.2019 **	-0.2181 **	-0.2259 **
nWhite	-0.0435 **	-0.0395 **	-0.0360 **	-0.0640 **
married	0.0764 **	0.0758 **	0.0738 **	0.0767 **
5 year diff of AgeXEdXOcc FTFY Empl	0.0108			0.0150 *
1 year diff of AgeXEdXOcc FTFY Empl	0.0027			-0.0041
5 year diff of AgeXEd FTFY Empl		0.0982 **		
1 year diff of AgeXEd FTFY Empl		-0.0577 †		
5 year diff of FTFY Ed Emp			0.1505 **	
1 year diff of FTFY Ed Emp			-0.0215	
Female				-0.0246 **
Constant	**	**	**	**
Observations	32080	32092	32112	41844

The marginal effects for our alternative definitions of lifetime occupation are:

Table 16: Marginal Effects of Probit Estimation- Dependent Variable=1 if Employed. Lifetime Occupation Spec. 2

	(1)	(2)	(3)	(4)
Wlimit1vAM	-0.1711 **	-0.1691 **	-0.1667 **	-0.2166 **
Wlimit2vAM	-0.5977 **	-0.5910 **	-0.5860 **	-0.6058 **
ageD2	-0.0216 **	-0.0188 **	-0.0171 **	-0.0043
ageD3	-0.1329 **	-0.1124 **	-0.1269 **	-0.1082 **
ageD4	-0.2133 **	-0.1852 **	-0.2013 **	-0.1991 **
nWhite	-0.0335 **	-0.0297 **	-0.0266 **	-0.0531 **
married	0.0525 **	0.0522 **	0.0503 **	0.0543 **
5 year diff of AgeXEdXOcc FTFY Empl	0.0127 †			0.0172 **
1 year diff of AgeXEdXOcc FTFY Empl	0.0062			0.0009
5 year diff of AgeXEd FTFY Empl		0.0867 **		
1 year diff of AgeXEd FTFY Empl		-0.0364		
5 year diff of FTFY Ed Emp			0.1390 **	
1 year diff of FTFY Ed Emp			-0.0232	
Female				-0.0364 **
Constant	**	**	**	**
Observations	27101	27112	27132	35509

Table 17: Marginal Effects of Probit Estimation- Dependent Variable=1 if Employed. Lifetime Occupation Spec. 3

	(1)	(2)	(3)	(4)
Wlimit1vAM	-0.2044 **	-0.2022 **	-0.1987 **	-0.2626 **
Wlimit2vAM	-0.6702 **	-0.6661 **	-0.6631 **	-0.6815 **
ageD2	-0.0296 **	-0.0267 **	-0.0260 **	-0.0165 *
ageD3	-0.1639 **	-0.1416 **	-0.1606 **	-0.1382 **
ageD4	-0.2649 **	-0.2387 **	-0.2569 **	-0.2408 **
nWhite	-0.0432 **	-0.0388 **	-0.0370 **	-0.0593 **
married	0.0726 **	0.0725 **	0.0710 **	0.0751 **
5 year diff of AgeXEdXOcc FTFY Empl	0.0063			0.0135
1 year diff of AgeXEdXOcc FTFY Empl	0.0098			0.0005
5 year diff of AgeXEd FTFY Empl		0.0822 **		
1 year diff of AgeXEd FTFY Empl		-0.0271		
5 year diff of FTFY Ed Emp			0.0891 **	
1 year diff of FTFY Ed Emp			0.1210	
Female				-0.0346 *
Constant	**	**	**	**
Observations	14985	14992	15004	20048

Table 18: Marginal Effects of Probit Estimation- Dependent Variable=1 if Employed. Most recent Occupation

	(1)	(2)	(3)	(4)
Wlimit1vAM	-0.1315 **	-0.1307 **	-0.1289 **	-0.1851 **
Wlimit2vAM	-0.6165 **	-0.6115 **	-0.6079 **	-0.6270 **
ageD2	-0.0170 **	-0.0148 **	-0.0147 **	-0.0204 **
ageD3	-0.1330 **	-0.1149 **	-0.1284 **	-0.1373 **
ageD4	-0.2399 **	-0.2147 **	-0.2308 **	-0.2527 **
nWhite	-0.0275 **	-0.0251 **	-0.0232 **	-0.0397 **
married	0.0414 **	0.0410 **	0.0396 **	0.0425 **
5 year diff of AgeXEdXOcc FTFY Empl	0.0059			0.0092 **
1 year diff of AgeXEdXOcc FTFY Empl	0.0061			-0.0026
5 year diff of AgeXEd FTFY Empl		0.0486 **		
1 year diff of AgeXEd FTFY Empl		-0.0414 *		
5 year diff of FTFY Ed Emp			0.0775 **	
1 year diff of FTFY Ed Emp			-0.0786 *	
Female				-0.0028
Constant	**	**	**	**
Observations	30252	30274	30296	39029

3.3 Wage Regression (With Heckman Selection Two-step)

We adjust for selection in the wage regression by implementing a two-step procedure following Heckman (1979). For the selection equation, we use the probit estimations above. We calculate the inverse Mills ratio from this equation and estimate a wage equation via ordinary least squares. The dependent variable is log hourly wage, excluding observations of more than $200 per hour or less than $3 per hour in CPI deflated 1999 US dollars.

Table 19: Wage Estimation- Dependent Variable Log Hourly Wage; Lifetime Occupation Specification 1

	(1)	(2)	(3)	(4)
Severe WL	-0.1624	-0.2661 **	-0.2629 **	-0.2001 *
	(0.1149)	(0.1014)	(0.0985)	(0.0833)
Moderate WL	-0.0688 *	-0.0969 **	-0.0953 **	-0.0816 **
	(0.0336)	(0.0301)	(0.0293)	(0.0290)
ageD2	-0.0320 **	-0.0324 **	-0.0339 **	-0.0222 *
	(0.0110)	(0.0107)	(0.0108)	(0.0094)
ageD3	-0.1033 **	-0.1200 **	-0.1232 **	-0.1049 **
	(0.0293)	(0.0262)	(0.0266)	(0.0220)
ageD4	-0.1470 **	-0.1739 **	-0.1744 **	-0.1600 **
	(0.0449)	(0.0406)	(0.0403)	(0.0352)
mills	0.1519	0.2548 **	0.2534 **	0.1878 *
	(0.1071)	(0.0949)	(0.0926)	(0.0778)
Onet Physical	-0.0394 **	-0.0394 **	-0.0395 **	-0.0259 *
	(0.0124)	(0.0124)	(0.0124)	(0.0101)
Onet KSA1	-0.0222 †	-0.0224 †	-0.0222 †	-0.0109
	(0.0119)	(0.0119)	(0.0119)	(0.0095)
Onet KSA2	0.0274 **	0.0274 **	0.0274 **	0.0257 **
	(0.0060)	(0.0060)	(0.0060)	(0.0051)
married	0.0494 **	0.0607 **	0.0600 **	0.0537 **
	(0.0179)	(0.0169)	(0.0166)	(0.0153)
time	0.0142 †	0.0185 *	0.0211 **	0.0181 **
	(0.0077)	(0.0079)	(0.0081)	(0.0068)
time2	-0.0021	-0.0028 *	-0.0032 *	-0.0023 *
	(0.0013)	(0.0013)	(0.0014)	(0.0011)
time3	0.0001 *	0.0002 **	0.0002 **	0.0001 *
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	2.7462 **	2.7207 **	2.7183 **	2.6470 **
	(0.0274)	(0.0270)	(0.0274)	(0.0222)
Observations	19052	19056	19064	24040

Standard errors in parentheses

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

The estimates for alternatives are:

Table 20: Wage Estimation- Dependent Variable Log Hourly Wage; Lifetime Occupation Specification 2

	(1)	(2)	(3)	(4)
Wlimit2vAM	-0.1175	-0.2781 **	-0.2707 **	-0.1774 *
	(0.1177)	(0.1044)	(0.1007)	(0.0825)
Wlimit1vAM	-0.0490	-0.0931 **	-0.0903 **	-0.0719 *
	(0.0347)	(0.0313)	(0.0303)	(0.0290)
ageD2	-0.0299 **	-0.0322 **	-0.0339 **	-0.0213 *
	(0.0115)	(0.0111)	(0.0112)	(0.0097)
ageD3	-0.0992 **	-0.1297 **	-0.1323 **	-0.1043 **
	(0.0313)	(0.0279)	(0.0282)	(0.0221)
ageD4	-0.1432 **	-0.1897 **	-0.1892 **	-0.1683 **
	(0.0470)	(0.0424)	(0.0420)	(0.0353)
mills	0.1199	0.2959 **	0.2901 **	0.1902 *
	(0.1214)	(0.1082)	(0.1047)	(0.0834)
Scores for component 1	-0.0269 †	-0.0274 †	-0.0273 †	-0.0133
	(0.0144)	(0.0144)	(0.0144)	(0.0117)
Scores for component 1	-0.0088	-0.0093	-0.0090	0.0030
	(0.0137)	(0.0137)	(0.0137)	(0.0110)
Scores for component 2	0.0226 **	0.0227 **	0.0227 **	0.0210 **
	(0.0069)	(0.0069)	(0.0069)	(0.0058)
married	0.0357 *	0.0496 **	0.0486 **	0.0418 **
	(0.0171)	(0.0164)	(0.0161)	(0.0151)
time	0.0169 *	0.0216 **	0.0248 **	0.0218 **
	(0.0081)	(0.0083)	(0.0086)	(0.0071)
time2	-0.0024 †	-0.0032 *	-0.0038 **	-0.0028 *
	(0.0013)	(0.0014)	(0.0014)	(0.0012)
time3	0.0002 *	0.0002 **	0.0002 **	0.0002 **
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	2.7859 **	2.7532 **	2.7504 **	2.6740 **
	(0.0264)	(0.0264)	(0.0269)	(0.0216)
Observations	16457	16461	16469	20921

Standard errors in parentheses

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

Table 21: Wage Estimation- Dependent Variable Log Hourly Wage; Lifetime Occupation Specification 3

	(1)	(2)	(3)	(4)
Wlimit2vAM	0.1803	0.0282	0.0321	-0.1095
	(0.1671)	(0.1557)	(0.1541)	(0.1192)
Wlimit1vAM	-0.0084	-0.0493	-0.0479	-0.0769 †
	(0.0478)	(0.0445)	(0.0444)	(0.0407)
ageD2	-0.0126	-0.0180	-0.0179	-0.0163
	(0.0157)	(0.0152)	(0.0154)	(0.0129)
ageD3	-0.0426	-0.0751 †	-0.0753 †	-0.0965 **
	(0.0431)	(0.0396)	(0.0411)	(0.0303)
ageD4	-0.0602	-0.1081 †	-0.1068 †	-0.1591 **
	(0.0641)	(0.0595)	(0.0604)	(0.0466)
mills	-0.1202	0.0213	0.0179	0.1368
	(0.1498)	(0.1401)	(0.1397)	(0.1032)
Scores for component 1	-0.0610 **	-0.0609 **	-0.0618 **	-0.0566 **
	(0.0187)	(0.0187)	(0.0187)	(0.0149)
Scores for component 1	-0.0357 *	-0.0357 *	-0.0363 *	-0.0355 *
	(0.0178)	(0.0178)	(0.0177)	(0.0139)
Scores for component 2	0.0274 **	0.0273 **	0.0278 **	0.0318 **
	(0.0092)	(0.0092)	(0.0092)	(0.0075)
married	0.0221	0.0370	0.0370	0.0482 *
	(0.0248)	(0.0241)	(0.0240)	(0.0212)
time	0.0230 *	0.0235 *	0.0235 *	0.0285 **
	(0.0109)	(0.0111)	(0.0111)	(0.0095)
time2	-0.0036 †	-0.0037 *	-0.0037 *	-0.0041 *
	(0.0018)	(0.0019)	(0.0019)	(0.0016)
time3	0.0002 *	0.0002 *	0.0002 *	0.0002 **
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	2.8390 **	2.8142 **	2.8145 **	2.6871 **
	(0.0362)	(0.0369)	(0.0364)	(0.0291)
Observations	9193	9196	9200	11793

Standard errors in parentheses

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

Table 22: Wage Estimation- Dependent Variable Log Hourly Wage; Most Recent Occupation

	(1)	(2)	(3)	(4)
Wlimit2vAM	-0.0875	-0.1524 *	-0.1385 †	-0.1354 *
	(0.0804)	(0.0773)	(0.0760)	(0.0629)
Wlimit1vAM	-0.0449 *	-0.0589 **	-0.0568 **	-0.0537 **
	(0.0215)	(0.0210)	(0.0208)	(0.0206)
ageD2	-0.0299 **	-0.0306 **	-0.0310 **	-0.0238 *
	(0.0107)	(0.0107)	(0.0107)	(0.0096)
ageD3	-0.0896 **	-0.1034 **	-0.1016 **	-0.0976 **
	(0.0252)	(0.0244)	(0.0246)	(0.0210)
ageD4	-0.1221 **	-0.1495 **	-0.1397 **	-0.1425 **
	(0.0395)	(0.0384)	(0.0379)	(0.0331)
mills	0.0800	0.1491 *	0.1349 †	0.1231 *
	(0.0777)	(0.0755)	(0.0745)	(0.0596)
Scores for component 1	-0.0437 **	-0.0434 **	-0.0447 **	-0.0311 **
	(0.0124)	(0.0124)	(0.0123)	(0.0101)
Scores for component 1	-0.0266 *	-0.0256 *	-0.0272 *	-0.0161 †
	(0.0120)	(0.0119)	(0.0118)	(0.0096)
Scores for component 2	0.0301 **	0.0306 **	0.0302 **	0.0282 **
	(0.0060)	(0.0060)	(0.0060)	(0.0050)
married	0.0394 **	0.0441 **	0.0429 **	0.0436 **
	(0.0142)	(0.0141)	(0.0141)	(0.0134)
time	0.0151 *	0.0171 *	0.0180 *	0.0189 **
	(0.0077)	(0.0077)	(0.0078)	(0.0068)
time2	-0.0022 †	-0.0025 †	-0.0026 *	-0.0024 *
	(0.0013)	(0.0013)	(0.0013)	(0.0011)
time3	0.0001 *	0.0002 *	0.0002 *	0.0001 *
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	2.7597 **	2.7501 **	2.7504 **	2.6665 **
	(0.0192)	(0.0196)	(0.0199)	(0.0160)
Observations	19246	19254	19265	24282

Standard errors in parentheses

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

3.4 Occupation Wage Trends.

We measure "structural decline" of an occupation as a fall in the wage payment to skills used to perform tasks comprising the occupation. We consider that individuals' stocks of skills are best suited for their life-time occupation. Therefore, we consider that their wages are related to the skill portfolio best matched to their life-time occupation even if we see them change occupations later in life. This raises the possibility for mis-match in later life changes; that an individual working say in construction for his whole life will be paid less if he switches to a service sector job than a comparable worker who has worked in services his whole life. Our goal then is to track changes in the wages paid to these skill-types over-time and then describe occupational wage changes as the change in the skill payments comprising that occupation.

The specific regression we consider is:

$$ln(w_{it}) = \beta^d \mathbf{X_{it}} + \beta^t \mathbf{T_t} + \beta^o \mathbf{O_i} + \beta^{ot} \mathbf{T_t} \times \mathbf{O_i}$$

The first regressor is a vector of demographic variables including a quadratic in experience, and dummy variables for each of: high school degree, non-white, and married. The second $\mathbf{T_t}$ is a quadratic in year with 1984 as the base, 1985 = 1 and so on.[7] The third $\mathbf{O_i}$ is a triple including the first principle component of the Onet physical tasks and the first and second principle component of the Onet knowledge-skill tasks in the individuals lifetime occupation.[8] The final term is an interaction of the time-quadratic with the Onet task triple. The sample selection is our base, excluding those with a college degree.

The full regression table is as follows.

[7] The quadratic trend provided the best representation of the data compared to cubic or time-year dummies. Estimates for these specifications including residual plots are available upon request.

[8] All results here are presented for our preferred lifetime occupation specification number 2.

Table 23: Wage Estimation- Dependent Variable Log Hourly Wage; Most Recent Occupation

	(1) Coefficient Estimate	Standard Error
Experience	.0127895	.0016915
Experience^2	-.0002885	.0000533
High School Degree	.2309461	.0082561
Non-White	-.0982968	.0082627
Married	.1079396	.011131
Occupation Physical	-.2152528	.0295344
Occupation Knowledge-Skill 1	-.1503618	.0293289
Occupation Knowledge-Skill 2	.1413814	.0154822
Time	-.0002879	.0022717
Physical $\times$ time	.0064967	.0045805
Knowledge-Skill 1 $\times$ time	.010728	.0045861
Knowledge-Skill 2 $\times$ time	-.0026212	.002354
Time^2	-.000032	.0000723
Physical $\times$ Time^2	-.000036	.0001429
Knowledge-Skill 1 $\times$ Time^2	-.0001274	.0001439
Knowledge-Skill 2 $\times$ Time^2	.0000111	.0000726
Constant	2.389939	.0204081
Observations	18144	
R-squared	0.1211	

The decomposition of occupational wages into the "price" paid to each task-skill along with the year trend components can be seen in Figure 6. It shows that the first principle component of Knowledge-Skill tasks have been a driver of wage growth. However, different occupations have different mixes of these components. The prediction for wage trends in each occupation based on how the price paid to tasks used in that occupation changes overtime can be seen in Figure 7. Occupations with declining payments to the tasks they use include household and building services, construction and extraction, production occupations, and most operator occupations.

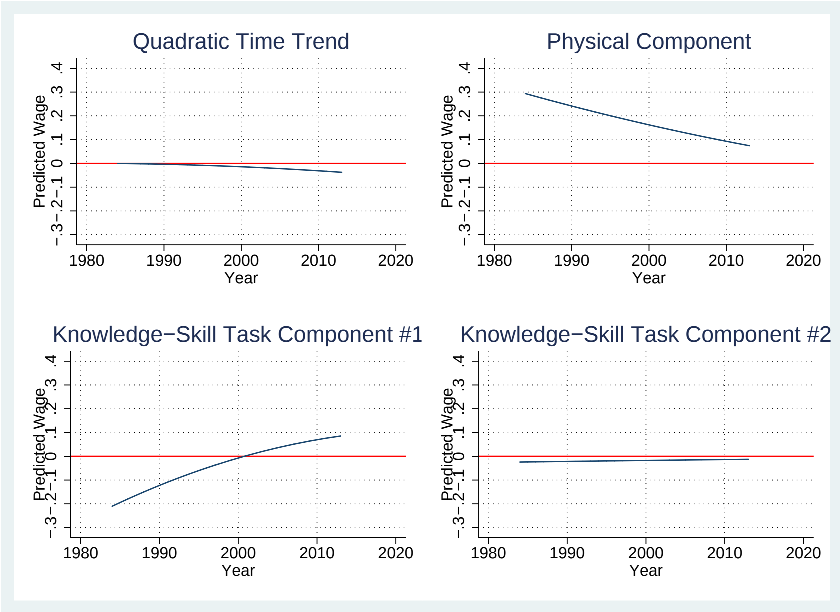


Figure 6: Predicted change in time and occupational task-skill component of wages.

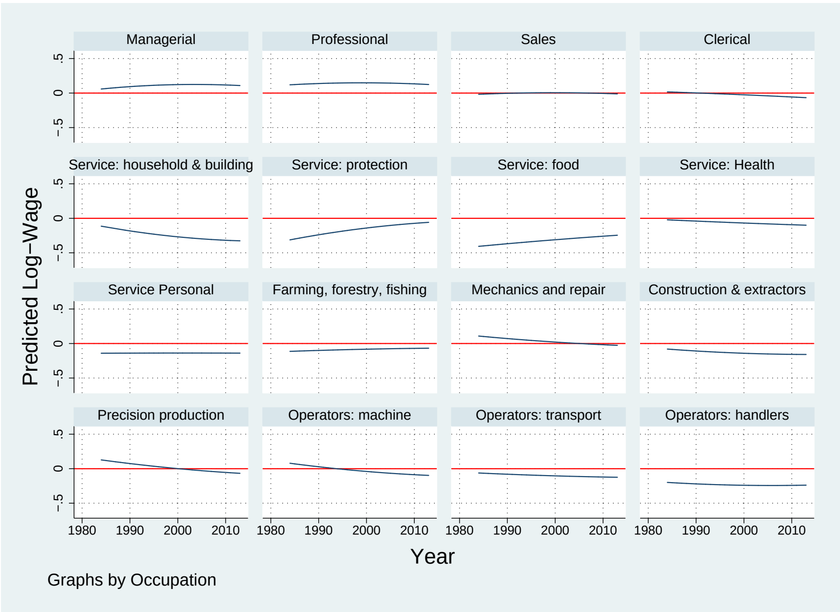


Figure 7: Predicted change in time and occupational task-skill component of wages.

Figure 8 provides a more concise definition of occupation. It groups the 16 SOC codes into quartiles of 4 occupations each according to their physical task intensity. Clearly, the most physically intensive occupations have suffered the largest predicted wage declines. This is important for our analysis because we have shown that the physical task intensity of an occupation is a strong predictor of both reported work limitations and disability receipt.

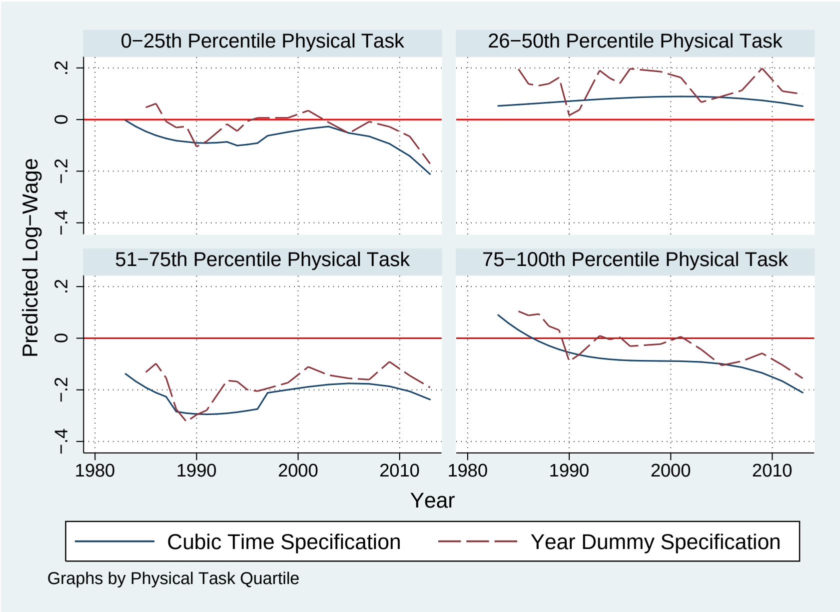


Figure 8: Predicted change in time and occupational task-skill component of wages.

Finally, we include Figure 9 to show our results are not driven by our cubic specification of time trends. The alternative specification in which we include year dummies and their interaction with the three task categories (long-dashed line) overlaps well with the cubic specification. The raw wages are also plotted. Be aware that the high-frequency variation in raw wages are an artifact of the small sample size of the PSID when divided into 16 occupation codes in addition to our baseline sample restrictions.

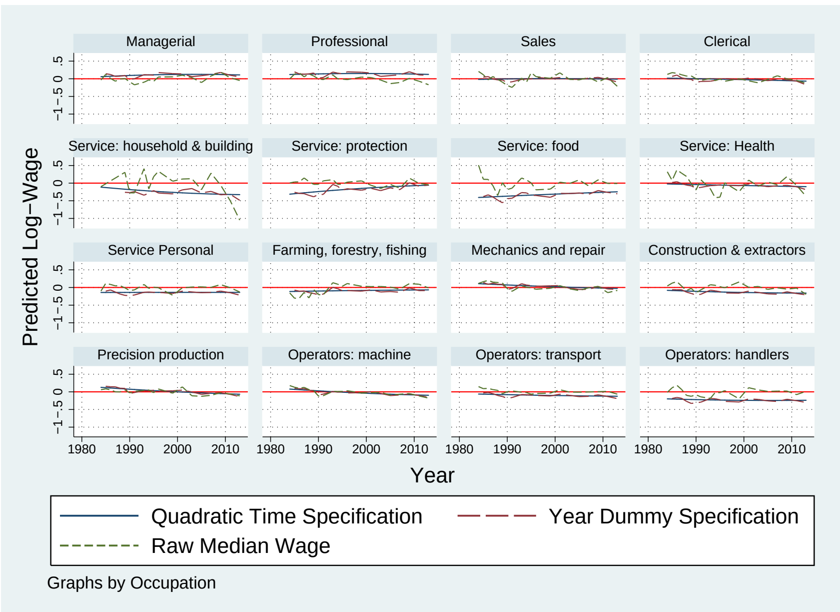


Figure 9: Predicted change in time and occupational task-skill component of wages and actual raw median wage.

4 Current Population Survey (CPS)

We use the current population survey to calculate changes in the full-time, full-year employment to population ratio for age-education demographic cells. We use this as the exclusion restriction in the selection equation in our two-step wage regression in the PSID. Our measure of cyclical risk in the model is also computed from the CPS. It is the job finding and job loss rates by occupation over the cycle.

Exclusion Restriction for Wage Probit. For our exclusion restriction we calculate the one and five year changes in full-time full year employment within occupation groups. The sample is limited to individuals not self-employed or in the military. Full-time, full-year is defined as usually more than 30 hours per week for greater than 49 weeks. Occupations are bridged from Census 1990 or 2000 codes (as applicable) to the 1980's codes and then to SOC codes by the scheme listed in Section 2. The sample cover the years 1980-2014. All data are weighted by the supplemental weights provided. Education categories are broken into three groups. They are (1) less than high school measured as less than grade 12 schooling ; (2) high school measured as completing at least grade 12 but not 4 years or more of college ; (3) four years or more of college.

Job Loss and Finding Rates. We use the month-to-month individual linking provided by Ruggles et al. (2013) and compute the fraction of employed workers who separate into unemployment and the number of unemployed workers who find a job. Separations are attributed to an occupation according to the main job in the month before unemployment while job finds are attributed to an occupation according to the occupation reported by the unemployed worker as their last occupation before unemployment. In sum, we associate unemployment transitions with the occupation from which the unemployment spell originated.

To convert these to cycle- and occupation-specific transition rates we use the time-aggregation correction from Elsby et al. (2009). Using NBER-defined recession dates, we take the average finding and separation rates for each occupation in recessions and expansions.

The specific time-series of the employment flow hazards to and from unemployment are shown by occupation in Figure 10. Figure 11 shows the mean flow rate and the standard deviation of the annual difference in flow rates.[9]

[9]Standardized statistic presented: $\hat{x} = \frac{x-\mu_x}{\sigma_x}$.

Figure 10: Time series of employment flows across occupations.

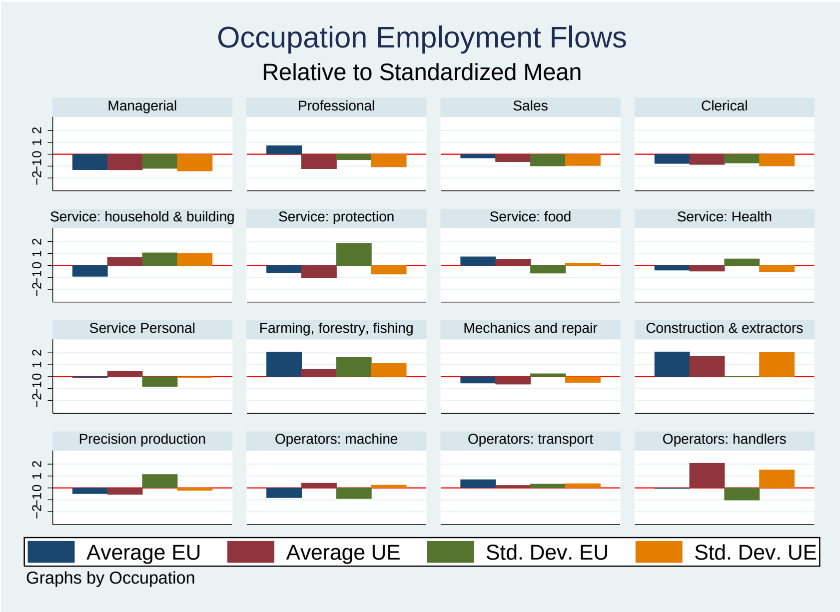


Figure 11: Variation in average and cyclical employment flows across occupations.

5 O*NET

We use the O*NET, a US Department of Labor database created to help workers understand the requirements of various occupations, to measure the task content of each occupation. This task content defines how occupations affect health outcomes. 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 2 principal components, slightly less than 70% of the variation. Finally, we de-mean and standardize each of the components. To merge O*NET occupation codes, to our coarser, 2-digit SOC codes, we take a simple average across occupations within the SOC categories. Figure 12 summarizes our findings by occupation.

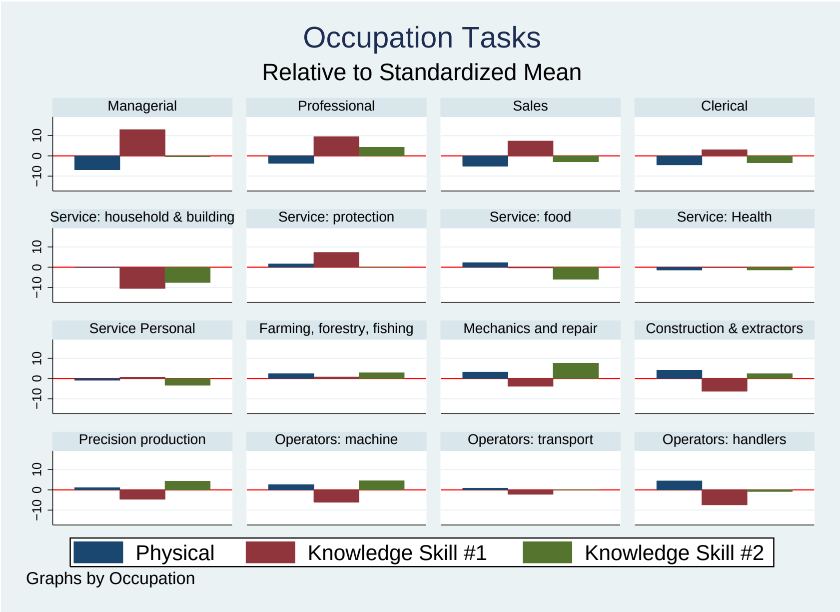


Figure 12: Variation in task intensity across occupations.

6 SSDI Application Acceptance Probability

We use the marginal effects from Lahiri et al. (1995) to construct the following two-step acceptance probability for an SSDI applicant conditional on their health, age, and the occupational productivity shock z_j in their occupation j. The first step is the probability an applicant is awarded benefits for severe enough health considerations $\pi_h(d, \tau)$. It depends on the extent of their work limitation d and their age τ. The second step is the probability an applicant is awarded benefits based on a combination of work limitations and vocational concerns, conditional on not being awarded on the first step: $\pi_v(\tau, z)$. It depends on their age τ and the occupational productivity shock z_j.[10] Therefore, the total award probability is: $\pi_{award}(d, \tau, z) = \pi_h(d, \tau) + (1 - \pi_h(d, \tau))\pi_v(\tau, z)$

The first step award probabilities are:

	None ($d = 0$)	Moderate ($d = 1$)	Severe ($d = 2$)
Age < 45	0.297	0.427	0.478
Age 45-55	0.315	0.450	0.508
Age 55-62	0.315	0.450	0.508

The regression in Lahiri et al. (1995) includes a dummy for one or more severe IADLs and a dummy for three or more severe ADLs. We assign the marginal effect of the former/latter to the moderate/severe limitation agents in our model $(d = 1)/(d = 2)$, respectively. The include only an age dummy for individuals less than 35, the marginal effect of which we assign to our youngest age group.

[10]See the main text for a discussion of how these factors are explicitly defined in the SSA rules and regulations.

The second step award base conditional probabilities are bounded between 17.14% and 39%, the second step conditional award probabilities observed in 1993 and 2010. We assume that this probability is linearly increasing in z_j. To this base, we add an additional acceptance probability of 12.4 percentage points for agents over the age of 55 to match the marginal effect of a corresponding dummy in Lahiri et al. (1995)

References

Elsby, M. W., R. Michaels, and G. Solon (2009): "The ins and outs of cyclical unemployment," *American Economic Journal: Macroeconomics*, 1, 84–110.

Lahiri, K., D. R. Vaughan, and B. Wixon (1995): "Modeling SSA's sequential disability determination process using matched SIPP data," *Soc. Sec. Bull.*, 58, 3.

Low, H., C. Meghir, and L. Pistaferri (2015): "Disability Insurance and the Dynamics of the Incentive-Insurance Tradeoff," *American Economic Review*.

Michaud, A. M. and D. Wiczer (2014): "Occupational hazards and social disability insurance," Working Papers 2014-24, Federal Reserve Bank of St. Louis.

Ruggles, S., J. T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek (2013): "Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]," Tech. rep., Minneapolis: University of Minnesota.