Data integration to quantify occupational injury risk wage premium

Paweł Strawiński, Dorota Celińska

University of Warsaw

2nd Congress of Polish Statistics,
11th July 2018

Financed by National Science Centre, Poland grant 2015/19/B/HS4/03231

Occupational injuries

- Occupational injuries have strong impact on the economy.
- According to the ILO almost 1 million of work-related accidents occur every day worldwide.
- The economic cost of those accidents is estimated on 4% of Global Gross Domestic Product each year (ILO, 2016).

Occupational risk

- *Occupational risk* is regarded as the possibility of a worker suffering a particular work-related injury.
- *Accident at work* is understood as a sudden event caused by external reason leading to injury or death which happened in connection with work.
- The are three distinct types of accident at work: fatal, serious and non-serious.

Popular approaches

- The standard treatment of occupational risk in the labour market is conducted in terms of the theory of compensating wage differentials.
- Those employed in risky job tasks that are associated with higher occupational injury risk are given wage premiums (positive relationship between wages and job risk in the market).
- There exists two strands of literature:
  1. Finding out determinants of sorting workers between safe and risky jobs;
  2. Calculation the implicit value of life or per statistical injury
- In both, dominant approach is to rely upon some published measure of the risk level by occupation at various degrees of aggregation, and then to match this risk variable with the data from the survey of workers

Cost of injury estimation

- Numerous studies, in particular for the US (Viscusi and Aldy, 2003) have shown that individuals in occupations with high injury risk are compensated for that risk by corresponding bonus payments.
- Male workers are overrepresented in the most dangerous occupations, while females typically work in relatively safe occupations with respect to occupational injuries.
- Workers sorting across occupations is also confirmed by many studies.
- Therefore, differences in occupational risk can partly explain the existence of the gender wage gap.

Occupational risk and wages

- Differences in occupational risk can partly explain the gender wage gap (King 1992; DeLeire and Levy 2004; Kuhn and Ruf 2009).
- Hersh (1998) showed that that women face a job risk that is about 70% that of men. Despite that, men and women receive similar wage compensation for occupational injury risk.
- Wage premiums can be different for different countries:
  - eg. in the USA nonfatal injury risk for men scores for 0.9-1.4% wage premium, while for women 2.8% -4.5% (Leeth and Ruser 2003),
  - in Switzerland there is negative or zero compensation for nonfatal accident risk at the workplace (Kuhn and Ruf 2009).
- Kluve and Schaffner (2007) results showed that workers are compensated for increased fatal occupational injury risk.

Motivation and aims

- Due to the data and methodological problems studies of influence of occupational risk on wages are rare.
- Most of the research relies upon some published aggregated measure of the risk level combined with data from the survey of workers.
- According to our best knowledge, there are no similar studies for Poland and few for other countries.
- We investigate at which level of data aggregation is possible to receive precise results.
- We are aiming at assessing size of wage premium with regard to the occupational risk in Poland.
- We also would like to investigate how the differences in the occupational injury risk of the jobs can help in explaining observed gender wage differentials.

Statistical card of accident at work

- A part of the programme of statistical surveys of the public statistic.
- Covers all accidents at work, as well as accident considered equivalent to accidents at work.
- As a single accident at work is counted each person's accident regardless whether the accident was an individual or a collective one.
- Database contains information on accidents, their effects, several personal and employers' characteristics.
- To have accident occurrence for possibly narrow group of workers we utilized the data from 2011-2014.

Number of accidents at work in Poland (10+workers)

- 95% of accident at work occur in firms employing 10+ workers

Gender	Severity	2011	2012	2013	2014
Women	non-serious	29,925	28,752	30,441	30,937
	serious	64	65	54	56
	fatal	14	21	14	10
	overall	30,003	28,838	30,509	31,003
Men	non-serious	60,981	55,981	51,988	51,916
	serious	544	465	399	398
	fatal	305	257	203	196
	overall	61,830	56,703	52,590	52,510
Total		91,833	85,541	83,099	83,513

Accidents by occupation, 2011-2014

men		
224 Paramedical practitioners	312.90	155
835 Ships' deck crew and related workers	65.05	319
612 Animal producers	55.70	2114
342 Sports and fitness workers	51.98	1462
532 Personal care workers in Health services	45.94	3265
511 Travel attendants, conductors and guides	37.26	4475
women		
612 Animal producers	41.67	1110
832 Car, Van and motorcycle drivers	39.11	326
833 Heavy truck and bus drivers	38.27	2182
511 Travel attendants, conductors and guides	30.92	3081
751 Food processing and related trade workers	26.61	45532
812 Metal processing and finishing plant operators	25.38	6038

Data aggregation

- We calculated accident ratios per 1000 workers for each gender $\times$ section $\times$ 4-digit occupation group (2x19x440).
- Unfortunately, we received large number of empty cells (about 50%).
- Moreover, for nearly 39% of the combinations of economic section, gender, and four-digit ISCO code in the accident database, we were unable to match observations from the SWS data on wages

Accidents measures

- We calculated accident ratios per 1000 workers for each gender $\times$ section $\times$ 3-digit occupation group (2x19x130). We eliminated groups containing 5 or less observations.
- Alternatively, we calculated accident ratio per 1000 workers for each gender $\times$ section group and for each gender $\times$ 3-digit occupation group. Then as an accident measure we used product of these two numbers divided by the average value of first measure.
- The alternative measure prevents from too many 0 accidents cells and high risk ratio for cells with few observations.

Data merging proces

- We merge data from the Statistical Card on Accident at Work database with those from the SWS survey.
- We use exact matching on economic sector (NACE section), three digit level of ISCO classification and gender.
- For each combination of those variables in the SWS database we link computed average risk ratios and our alternative measure.
- We match information for 4134 out of 4175 distinct groups (over 99%). This is due the fact that SWS is a representative survey.

Mincer's equation

- We start with simple hedonic wage regression, so called Mincer-type equation, of the following form:

$$ln(w_i)=\beta_0+\sum_{i=1}^{I}\beta_i PER_i+\sum_{j=1}^{J}\gamma_j JOB_i+\sum_{k=1}^{K}\delta_k ACC_k+\varepsilon_i$$

- w_i – gross monthly salary (PLN);
- PER_i – personal characteristics: highest level of formal education; number of years of work experience and it's square; gender of the employee;
- JOB_k – job characteristics: size of the firm (small, medium, big, very big), ownership sector (1 for public, 0 for private);
- ACC_k – accident ratios and interactions between them and gender.

Wage equation variants

1. full sample,
2. sample limited to workers 25-55 years old,
3. sample limited to full-time employed,
4. sample limited to full-time employed aged 25-55,
5. sample limited to full-time employed, groups with accident ratio above three times average excluded
6. sample limited to full-time employed aged 25-55, groups with accident ratio above three times average excluded
7. sample limited to full-time employed, three accident measures
8. sample limited to full-time employed, groups with accident ratio above three times average excluded, three accident measures

Results of estimation

Dependent variable: log monthly gross salary

	(1)	(2)	(3)	(4)	(5)	(6)
exp	.0260*	.0298*	.0255*	.0291*	.0255*	.0291*
exp_sq	-.0004*	-.0005*	-.0004*	-.0005*	-.0004*	-.0005*
female	-.2014*	-.2160*	-.2177*	-.2329*	-.2245*	-.2385*
acc	-.0025*	-.0028*	-.0027*	-.0030*	-.0078*	-.0081*
fem x acc	-.0007*	-.0008*	-.0004	-.0003	-.0018*	-.0018*
N	730,483	597,201	674,613	553,422	662,978	544,530
R2	.4259	.4163	.4336	.4244	.4401	.4307

* denotes significance at 1% level

- Results for an alternative measure are qualitatively the same

Results for separate severity indicators

Dependent variable: log monthly gross salary

	(7)	(8)
exp	.0255*	.0255*
exp_sq	-.0004*	-.0004*
female	-.2174*	-.2250*
acc_non	-.0031*	-.0091*
femXacc_n	-.0003	-.0010*
acc_ser	.0504*	.1572*
femXacc_s	.0988*	.0371
acc_fat	.0083	.0020
femXacc_f	-.0183	.0184
N	674,613	662,978
R2	.4338	.4406

* denotes significance at 1% level

O-B decomposition variants

1. total accidents, sample limited to full-time employed,
2. 3 severity types of accidents, sample limited to full-time employed,
3. total accidents, sample limited to full-time employed aged 25-55, groups with accident ratio above three times average excluded
4. 3 severity types of accidents, sample limited to full-time employed aged 25-55, groups with accident ratio above three times average excluded
5. total accidents alternative measure, sample limited to full-time employed,
6. total accidents alternative measure, sample limited to full-time employed aged 25-55, groups with accident ratio above three times average excluded

O-B decomposition results

Dependent variable: log monthly gross salary

	(1)	(2)	(3)	(4)	(5)	(6)
Difference	.1234*	.1234*	.1294*	.1294*	.1234*	.1432*
Explained	-.0977*	-.0945*	-.1108*	-.1023*	-.1273*	-.1240*
acc_rate	-.0180*	-.0134*	-.0457*	-.0337*	-.0611*	-.0627*
acc_rate (%)	.1838*	.1418*	.4125*	.3297*	.4795*	.5055*
Unexplained	.2211*	.2179*	.2401*	.2317*	.2507*	.2672*

* denote significance at 1% level.
In model (2) and (4) coefficient for fatal accident rate are not significant.

Conclusions

- Results for models with aggregate accident indicators suggest that workers are not properly compensated in wages for working in hazardous conditions.
- Possible explanations for observed phenomenon:
  - Measurement problem, accidents are rare events,
  - Underestimation of work related risk by employees,
  - Firms have more bargaining power in wage negotiations.

Conclusions

- More intuitive results are obtained from models with disaggregated accident measures
- Compensation for non-serious accident risk is statistically significant and negative.
- Compensation for serious accident risk is positive and estimated on 5% to 15%.
- Fatal accident risk is statistically insignificant – fatal accidents are (fortunately) rare.
- After Oaxaca-Blinder decomposition, the estimated wage gap is 0.13 with (-0.09 to -0.13) explained and (0.21 to 0.26) unexplained.
- The accident risk explains 15% to 30% of the explained gap, more than for instance work experience.