Does the Value per Statistical Life Vary with Age or Baseline Health? Evidence from a compensating wage study in France

Daniel Herrera-Araujo*, Lise Rochaix†

Abstract

This paper provides an empirical assessment of the effects of age and baseline health on the Value per Statistical Life (VSL) by reporting the results of a compensating wage differential for occupational fatality risk in France. We exploit Constances, a novel population-based cohort that combines respondents' full medical history, elicited using face-to-face interviews with physicians, with respondents' actual work history, extracted from administrative records. Focusing on blue-collar males, aged between 20 to 59 years of age, we find an average VSL estimate of 6.5 million euros. Our results support the hypothesis that VSL varies with age and baseline health: VSL decreases with age and weakly increases with baseline health.

JEL Codes: J3, J28, J17, I1.

Keywords: Value per Statistical Life, France, Wage Compensating Differentials, Age, Baseline Health.

*Corresponding author (e-mail: daniel.herrera@dauphine.psl.eu), Université Paris-Dauphine, PSL Research University, LEDa (CGEMP) UMR CNRS [8007], Paris, France.

†Hospinnomics, Paris School of Economics - University of Paris 1 Panthéon-Sorbonne
Acknowledgments: Special thanks to Marie Zins, Marcel Goldberg, Lisa Berkman, Julie Rochut, and the entire Constances team for useful comments on Constances and CNAV datasets. Thanks to Pascale Lengagne, Nathalie Serres and Pascal Jacquetin for useful comments and discussions on the CNAMTS dataset. Thanks also to Christoph Rheinberger, Matthew Adler, James K. Hammitt, Henrik Andersson, Laura Taylor, two anonymous referees, and participants at EAERE (2019) and at WONV (2019) for helpful comments and discussions. We declare that we have no conflict of interest. All errors are our own.

1 Introduction

All humans face a wide array of risks to health and life induced by environmental and other factors, but they differ in their valuations of wealth, health and life. It seems reasonable to expect that individuals' valuations for reducing risks to health and life depend on personal characteristics. That is, for a similar life- or health-risk reduction, the gains may not be valued in the same way by young (healthy) individuals compared to older (less healthy) ones. Thus, the willingness to pay (WTP) for a small mortality risk reduction, namely the Value per Statistical Life (VSL), may depend on the context in which it is valued. By combining a data set of French industry-related mortality risks, along with a population-based panel containing individuals' medical and work history, we provide the first revealed preference study (RP) investigating the relationship between VSL, age and baseline health.

As VSL is a key statistic used to measure the benefits of mortality risk reductions generated by public policies, it is important to both researchers and policy-makers across many fields of application. Yet, adopting a context-dependent VSL estimate for benefits' valuation may prove highly controversial. For instance, the US Environmental Protection Agency's (EPA) decision to use an age-adjusted VSL was strongly criticized by members of the general public,[1] and the agency quickly withdrew its proposal. Similarly, in their guidance on impact assessment, the European Commission and the UK Treasury recommended using context-specific VSL estimates, with higher estimates for a particular health condition (cancer), because of the devastating nature of the disease (European Commission 2001; HM Treasury 2003); a decision that has not generated the same public attention as in the US EPA's case.[2]

Insights from theory, however, suggest that VSL does vary with age and baseline health. Yet, there is no agreement as to how VSL varies with respect to each dimension. The signs and magnitude of the effects of age and baseline health on VSL are ambiguous. Some models predict that VSL should decline steadily with age (Jones-Lee, 1989), while others suggest that the age/VSL relationship could either be increasing, decreasing or even independent of age (Johansson, 2002; Aldy and Viscusi, 2004; and Ehrlich and Yin, 2005). Theoretical models linking VSL with health also predict ambiguous results (Hammitt, 2002; Bleichrodt et al., 2006; Rheinberger et al., 2016). In this paper, we define a simple theoretical framework that combines quality-of-life (defined here as a single index function which combines individuals' longevity with their health) and wealth dimensions. In line with previous theoretical models, we find ambiguous results. In the absence of clear theoretical predictions, the relationship between VSL, age and baseline health therefore remains an empirical question.

From an empirical perspective, most of the literature focusing on the link between age,

[1] The EPA, under the Clear Skies initiative, proposed to use a VSL estimate for those aged 65 and older that was 37% lower than for those aged 18-64 (Viscusi & Aldy, 2007).

[2] For the EU report, see https://ec.europa.eu/environment/enveco/others/pdf/recommended_interim_values.pdf [last visited on 30^{th} July 2019].

health and VSL is based on stated preferences (SP). An asset of these methods is the ability to combine health-related information with individuals' trade-off between wealth and mortality risks. SP methods consist of presenting individuals with hypothetical choices. They include contingent-valuation methods and discrete-choice experiments (Bateman et al., 2002). Alberini et al. (2004), provide an empirical assessment of the effects of age and baseline health on WTP for mortality risk reductions by reporting the results of two contingent valuation surveys. They find weak support for the hypothesis that WTP declines with age, cardio-vascular disease, lung disease, and cancer. Cameron and Deshazo (2013) estimate a structural model with survey data, allowing them to infer how WTP estimates for reductions in the risks of sick-years and lost life-years depend upon the individual's age, income, marginal utility of other consumption, and discount rate. They find a positive marginal utility associated with income and a negative marginal utility associated with the logarithms of present discounted sick-years, and lost life-years.

Using a RP method, there is a large body of literature assessing the relationship between VSL and age. For example, Viscusi and Aldy (2003) use an age-dependent fatality risk measure to estimate age-specific hedonic wage regressions. They find that VSL exhibits an inverted-U-shaped relationship with age. Hedonic wage studies generally support a flat or inverted U-shaped age-VSL relationship (O'Brien, 2018).[3] However, there is some variation. Evans and Schaur (2010) propose a quantile regression approach to provide insights into the relevance of earnings and age variation for VSL estimates. They find a decreasing relationship between age and VSL.

In contrast, to the best of our knowledge, the literature based on RP methods to value health related-risks in a VSL context is thin. Gayer et al. (2002) examine the effect on housing prices of cancer risks from chemical exposures from hazardous waste sites. In a similar study, Davis (2004) exploits a natural experiment to assess the marginal WTP to avoid a statistical case of pediatric leukemia risk. Gentry and Viscusi (2016) propose a methodological framework to distinguish the fatality and morbidity components of VSL. They show that VSL can be separated into two additive components: the morbidity associated with fatal injuries and the fatality per-se, providing the first RP estimates distinguishing both effects.

This article extends the existing literature by examining the effects of baseline health and age on VSL. We use individual level data on demographic characteristics from Constances, a large population-based cohort in France. It is a nationally representative sample of adults, which includes information on age, sex, medical history, and professional career. Constances matches detailed individuals' characteristics with their professional history. Annual information is collected on gross wages, the number of equivalent quarters worked per year and occupation, as well as firm and industry identifiers. For our analysis, we follow the VSL literature and focus on blue-collar males, aged between 20 to 59 years of

[3]Other papers investigating the relationship between age and VSL using a hedonic approach include Smith et al. (2004), Kiesner et al. (2006) and Aldy and Viscusi (2008).

age. Using industry identifiers, we link the Constances cohort with data on work-related mortality risks from the French National Health Insurance Fund. This publicly available data gathers information on accidents from all French workers who are covered by the National Technical Committee (NTC). It contains information on the total number of full-time equivalent workers (FTE), the number of hours worked, and the number of deaths on an annual basis per industry.[4]

To identify the effects of age and baseline health on VSL, two complementary approaches are used: the first uses cross-sectional variations, while the second uses panel variations. The cross-sectional analysis has a comparative advantage over the second in that Constances' wide range of variables allows us to control for a large range of potentially confounding factors that would otherwise lead to biased estimates. This large array of controls, however, is not available through time and cannot be used in the panel setting. On the other hand, panel data allows for more observations, while still controlling for unobserved time-invariant individual characteristics. The identification strategy exploits individual-level health variations, defined as the age of onset of either cardio-vascular disease or cancer, along with job changes, to identify the effect that baseline health may have on VSL.

Our results suggest that the average VSL estimate is close to 6.5 million euros (in 2015 euros), which is similar to recent RP-based estimates from fatalities at work.[5] Early compensating differential studies on fatalities at work found weak evidence of a risk premium for riskier jobs (Fairris, 1989; Dorman & Hagstrom, 1998). More recently, Viscusi (2004), using the U.S. BLS Census of Fatal Occupational Injuries (CFOI) finds a VSL of 7.5 million dollars (in 2000 dollars), which corresponds to 9.4 million euros (in 2015 euros). Using the U.S. Census Bureau's Census of Manufacturers along with a quasi-experimental setting based on random safety inspections, Lee & Taylor (2019) find VSL values between 8 and 10 million dollars (in 2016 dollars), which corresponds to 7.1 to 8.9 million euros (in 2015 euros).

We find that individuals with a cardio-vascular disease consistently require more compensation than healthy individuals for a similar mortality risk exposure. However, we find weak evidence that individuals with cancer would require a larger compensation. Individuals with cancer who are less than 50 years of age require greater compensation than healthy individuals for a similar exposure to mortality risk, while those older than 50 years require less compensation. Moreover, our results suggest that VSL varies with age. VSL estimates range from 13 million euros for individuals below 30, to 4 million for those

[4]The dataset covers all French industries per year. However, while it includes blue-collar fatality yearly counts on a four-digit level (720 industries), the total number of blue-collar workers per year is only available at the two-digit level (38 industries). The number of hours worked per industry, however, is only available from 2013.

[5]The VSL is based on disposable income, which is approximated by multiplying the VSL by a factor of 0.77. The VSL, which is computed using gross wages, equals 8.5 million euros (in 2015 euros).

above 30 (in 2015 euros). Our study, however, suffers from the same limitations as other hedonic wage studies. In particular, it does not capture preferences of those individuals who remain outside the study sample.

The structure of this paper is as follows. In the next section, we develop a theoretical model on the relationship between VSL and age/baseline health. In Section 3, we describe the data sets used for the empirical analyses. In Sections 4 and 5, we describe the empirical implementation and the results, respectively. A final section summarizes and concludes.

2 The model

Consider an individual who derives utility $u(w, q)$ from wealth w and quality-of-life q. Let quality-of-life be a single index function, which combines individuals' longevity with their health. This implies that any improvement in health or longevity improves quality-of-life. In the following, we denote first (second) derivatives of the utility function with respect to wealth by the subscript 1 (11) and those with respect to quality-of-life by the subscript 2 (22). Further, we assume non-satiation with respect to income: $u_1(w, q) > 0$; non-satiation with respect to quality-of-life: $u_2(w, q) > 0$; weak financial risk aversion: $u_{11}(w, q) \leq 0$, which states that less risk over wealth is preferable to more risk; and correlation affinity: $u_{12}(w, q) \geq 0$. This last assumption implies that the marginal utility of wealth does not decrease with better quality-of-life (Viscusi and Evans, 1990; Sloan et al.,1998; Finkelstein et al., 2013).

Consider now an individual who is facing a work-related risk resulting in one of two states of the world: in the next working period, she will either live or die. Let π denote the probability of dying; hence the survival probability is $1 - \pi$. Conditional on survival, quality-of-life is q. If the individual does not survive, quality-of-life is equivalent to that of being dead, and is denoted as $\underline{q}$.

The individual's expected utility is given by:

$$EU(w, q) = (1 - \pi)u(w, q) + \pi u\left(w, \underline{q}\right),$$

where $u(w, q)$ and $u\left(w, \underline{q}\right)$ are the utilities associated with wealth w conditional on the state of the world, with either baseline health or equivalent-to-death health, respectively.

Assume the individual is offered an opportunity to decrease the mortality probability, π, by the amount θ_π. In return for the decreased risk of premature death, the individual is willing to decrease current wealth by $C(w, h, \pi, \theta_\pi)$. By definition, this equals the amount that leaves the individual with the same expected utility as in the initial (pre-intervention) situation. Formally, the compensating variation is defined as:

$$\begin{aligned}(1-\pi^*)u\left(w - C\left(w,q,\pi,\theta_\pi\right),q\right) \\ +\pi^* u\left(w - C\left(w,q,\pi,\theta_\pi\right),\underline{q}\right) = EU(w,q),\end{aligned} \tag{1}$$

with $\pi^* \equiv \pi - \theta_\pi$. To simplify notation in the analysis presented below, let:

$$\begin{aligned} C\left(w,q,\pi,\theta_\pi\right) &\equiv C_\pi, \\ w - C_\pi &\equiv w^*, \\ (1-\pi^*)u_1\left(w^*,q\right) + \pi^* u_1\left(w^*,\underline{q}\right) &\equiv EU_1(w^*,q), \\ (1-\pi)u_1\left(w,q\right) + \pi u_1\left(w,\underline{q}\right) &\equiv EU_1(w,q), \\ (1-\pi)u_{12}\left(w,q\right) + \pi u_{12}\left(w,\underline{q}\right) &\equiv EU_{12}(w,q). \end{aligned}$$

2.1 Deriving values for fatality risk reductions

We obtain the corresponding marginal willingness to pay (MWTP) for a reduction in mortality risk by differentiating equation (1) with respect to θ_π:

$$MWTP_{\theta_\pi} \equiv \frac{\partial C_\pi}{\partial \theta_\pi} = \frac{u(w^*,q) - u(w^*,\underline{q})}{EU_1(w^*,q)} > 0, \tag{2}$$

where the numerator equals the gain in utility from the reduced risk of dying and the denominator represents the expected marginal utility of consumption. When the mortality-risk reduction approaches zero (i.e., $\theta_\pi = 0$), we find that:

$$MWTP_{\theta_\pi}\Big|_{\theta_\pi=0} \equiv \frac{\partial C_\pi}{\partial \theta_\pi}\Big|_{\theta_\pi=0} = \frac{u(w,q) - u(w,\underline{q})}{EU_1(w,q)} = VSL(q), \tag{3}$$

where $VSL(q)$ stands for the Value per Statistical Life at the quality-of-life q. As both expected marginal utility gains and costs are positive, a marginal reduction in mortality risk is valuable.

A better quality-of-life state, because it improves both the numerator and the denominator of equation (3), may increase or decrease VSL. On the one hand, surviving with better quality-of-life is more desirable. On the other hand, better quality-of-life increases the marginal utility of consumption, thus the opportunity cost of reducing mortality risks. The relationship between quality-of-life and VSL is obtained by differentiating equation (3) with respect to quality-of-life, q, as follows:

$$\frac{\partial VSL(q)}{\partial q} = \frac{u_2(w,q)}{EU_1(w,q)} - VSL(q)\frac{EU_{12}(w,q)}{EU_1(w,q)}. \tag{4}$$

The first term on the right-hand side (RHS) corresponds to the marginal gain from improving the individual's quality-of-life state. As individuals care about quality-of-life, this first term is positive. The second term on the RHS captures the effect of an improved quality-of-life state on the opportunity cost of spending resources on mortality risk reductions. Due to the correlation affinity between quality-of-life and wealth, this second term might also be positive. Thus, the effect of quality-of-life on VSL may be undetermined; hence, how VSL varies with age and health is an empirical question.

3 Data

The following section describes the two data sets used in the estimation: first, an individual level data set containing socio-economic characteristics, including medical and professional history; and second, an industry-wide data set that characterizes the occupational risks.

3.1 Individual level: Constances and CNAV data

We obtain individual level data on demographic characteristics from Constances, a large population-based epidemiological cohort. It is a nationally representative sample of adults and includes information on age, sex, health status, medical, and professional history. Information is collected through questionnaires sent to respondents' homes. Respondents, however, are also requested to come to a medical center for an exhaustive range of medical checks. Most of the health-related questions, including medical history, are elicited by physicians working at these centers.[6]

Constances matches individuals' detailed personal characteristics with their professional history extracted from the National Retirement Insurance Fund, administered by the French National Insurance Fund for the Elderly, hereafter CNAV.[7] This data set contains annual information on gross wages, the number of quarters validated in a year, as well as occupation, firms and industry identifiers of where respondents have been employed. For our analysis, we focus on blue-collar males aged between 20 and 59 years of age. The final data set contains the working histories for a panel of 7,268 blue-collar male workers between 2002 and 2016.[8,9]

[6]All participants in the sample live in, or close to, one of 20 cities/departments in France: Angoulême, Auxerre, Bordeaux, Caen, Haut-Rhin, Le Mans, Lille, Lyon, Marseille, Nancy, Nimes, Orléans, Paris, Pau, Poitiers, Rennes, Saint-Brieuc, Saint-Nazaire, Toulouse, Tours.

[7]In French it is called "Caisse Nationale d'Assurance Vieillesse" or CNAV. The system allows the collection of social and occupational data from different funds that manage various insurance schemes and other social transfers.

[8]Our data on working histories is available before 2002. In principle, we can use information from prior periods. However, we limit the sample to 2002 for two reasons: first, we only have national level information on mortality risk at work prior to 2009; and second, we observe a decreasing trend of the fatality rate at work in periods prior to 2002. Figure 2 depicts the evolution of fatality rates at work from 1955 up to 2016.

[9]To identify blue-collar workers, we exploit individuals' self-reported blue-collar status from the

Table 1: Summary statistics on Constances' selected variables

	Mean	Std. Dev.	Min	Max	Type
Average age	40.46	10.16	20.00	59.00	P
Less than 30 years old	0.18	0.38	0.00	1.00	P
Between 30 and 49 years old	0.60	0.49	0.00	1.00	P
More than 50 years old	0.22	0.42	0.00	1.00	P
Average gross hourly wage (in 2015 €)	14.59	5.43	1.25	29.65	P
Less than 8 years of education	0.04	0.18	0.00	1.00	C
Between 9 and 15 years of education	0.87	0.33	0.00	1.00	C
More than 16 years of education	0.09	0.29	0.00	1.00	C
Has a partner?	0.72	0.45	0.00	1.00	C
Considers himself to be happy	0.76	0.43	0.00	1.00	C
Considers himself as a smoker	0.31	0.46	0.00	1.00	C
Audit drinking score	5.75	4.76	0.00	38.00	C

Notes: The sample is limited to male blue-collar men aged between 20 and 59 years. The average hourly gross wage is computed by dividing the annual gross wage by 1,607, the legal maximum number of hours that an individual can work in a year. Respondents were asked to self-assess their smoking status: if they considered themselves to be smokers, the variable takes a value of 1 and 0 otherwise. AUDIT score ranges from 0 to 38. A score between 0 to 7 is considered as low-risk of dependence, 8 - 15 as hazardous levels of dependence, 16 - 20 as a high risk of dependence, and 20+ as almost certainly dependent.

Table 1 presents the summary statistics of the sample of respondents included in the analysis. The sample is composed of respondents aged between 20 and 59, with an average age of 40 years. Around 18% of respondents are below 30, 60% of respondents are between 30 and 49 years and the rest are above 50. The hourly gross wage is computed by dividing the annual gross wage by 1,607, the legal maximum number of hours that an individual can work in a year.[10] On average, an individual in our sample earns 14.59 euros per

survey part of Constances, and assume that this status was the same in previous years. Although CNAV reports information on blue-collar worker status, there are two limitations in using this variable: first, prior to 2012, the variable is rarely reported; and second, the blue-collar status is reported by firms, not employees. As a robustness check, we report the results using the CNAV blue collar variable. The coefficients are identified using on average four years per respondent, rather than nine years. The fatality rate coefficient remains positive and statistically significant.

[10]see https://www.service-public.fr/particuliers/vosdroits/F1911 [last visited on July 1^{st} 2019]

hour.[11,12] The proportion of respondents declaring between nine to 15 years of education is 87%, 4% declared having less than nine years, and 9% declared having more than 15 years. Most respondents report having a partner, and consider themselves to be happy most of the time.

In terms of lifestyle, we selected two questions one on smoking and the other on alcohol. A third of respondents (= 31%) considered themselves to be a smoker. Alcohol status was derived using an Alcohol Use Disorders Identification Test (AUDIT). The AUDIT is a 10-item screening tool developed by the World Health Organization (WHO) to assess alcohol consumption, drinking behavior and alcohol-related problems. A score between 0 to 7 is considered as low-risk of dependence, 8 - 15 as hazardous levels of dependence, 16 - 20 as a high risk of dependence, and 20+ as almost certainly dependent. As the AUDIT score average for our sample of respondents is 5.75 (on a scale from 0 to 38), the average respondent in our sample is considered to be low-risk.[13]

Although we observe respondents' professional history, including wages, we only have demographic characteristics for those years when respondents were surveyed. In order to use the panel dimension of Constances, we are restricted to using the demographic variables available that can be exploited with the panel dimension of our professional history data. While this is the case for some of the respondents' personal characteristics, it does not apply for others. For this reason, we categorize demographic variables into two groups: cross-section and panel variables. Variables in Table 1 of type "C" are considered for cross-sectional use only, while variables of type "P" are considered for both cross-sectional and panel use.

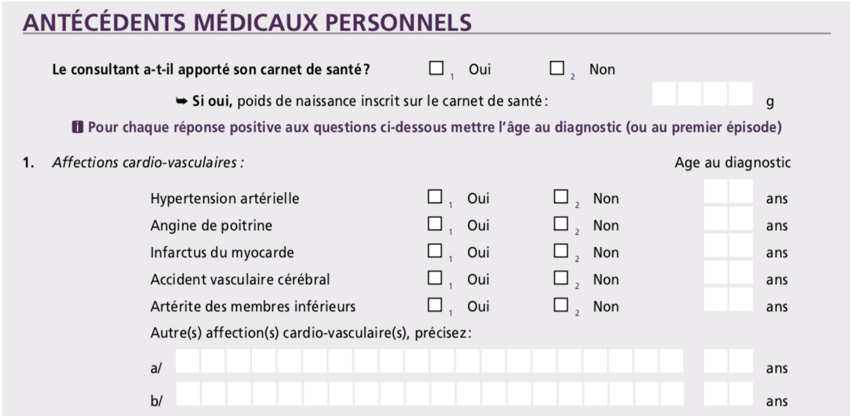
All variables about respondents' medical history are considered as type "P". As illustrated in Figure 1, respondents were asked by physicians whether they had been diagnosed with cardio-vascular disease or cancer.[14] Conditional on declaring the disease, respondents

[11] One of the limitations of the CNAV wage variable is that firms are obliged to report individuals' gross earnings up to a legally defined threshold. For any annual wage over the legal threshold, firms may decide whether or not to report or conceal the amount above the threshold. As a result, the annual gross wage may be truncated above the legal threshold, which, in addition, varies from year to year. Fortunately, this is not a major concern here as 93% of the wages are below the threshold. Moreover, if we exclude those wages above the threshold, average hourly wages drop to 13.86 (in 2015 euros) and, more importantly, our results are quantitatively and qualitatively the same. Information about the thresholds can be found here: https://www.legislation.cnav.fr/Pages/bareme.aspx?Nom=salaire_plafond_soumis_cotisation_bar [last visited on July 18th 2019].

[12] In France, the average percentage difference between gross and net is 23%. Estimates are adjusted by this difference and re-expressed in net terms. As a base for comparison, the national average wage for blue-collar workers in France was 13.36 € in 2015. Constances' wage data yields an average wage close to the national average. All wages are converted into 2015 euros using French CPI value available on the OECD website [last visited on 9th July 2019, https://data.oecd.org/price/inflation-cpi.htm].

[13] The AUDIT has been validated across genders and in a wide range of racial ethnic groups and is well suited for use in primary care settings (WHO).

[14] While other diseases, such as respiratory disease or digestive-related disease, are reported in the data, we chose to focus on cardio-vascular disease and cancer. Figure 1 presents a summary of



ANTÉCÉDENTS MÉDICAUX PERSONNELS

Le consultant a-t-il apporté son carnet de santé ? ☐ 1 Oui ☐ 2 Non

➥ **Si oui,** poids de naissance inscrit sur le carnet de santé : g

Pour chaque réponse positive aux questions ci-dessous mettre l'âge au diagnostic (ou au premier épisode)

1. *Affections cardio-vasculaires :*

			Age au diagnostic
Hypertension artérielle	☐ 1 Oui	☐ 2 Non	ans
Angine de poitrine	☐ 1 Oui	☐ 2 Non	ans
Infarctus du myocarde	☐ 1 Oui	☐ 2 Non	ans
Accident vasculaire cérébral	☐ 1 Oui	☐ 2 Non	ans
Artérite des membres inférieurs	☐ 1 Oui	☐ 2 Non	ans

Autre(s) affection(s) cardio-vasculaire(s), précisez:

a/ ans

b/ ans

Figure 1: Extract from the medical history questionnaire

were asked their age at the disease onset. As reported in Table 2, during our period of analysis, 12.64% of our sample report having suffered from cancer or cardio-vascular disease: 1.71% report cancer, and 10.92% report cardio-vascular disease. The average number of years for which we have information per respondent is close to 10 years. Healthy individuals report a duration of 9.69 years, while individuals affected by a disease report around 9.45 years. As compared with healthy individuals, individuals with a medical condition are older. At cancer onset, individuals are aged 46 years on average, while for cardio-vascular diseases individuals are aged 45 years at onset. Finally, individuals with a cancer or cardio-vascular disease change industry less often than healthy individuals. Those with neither condition experience, on average, 0.11 new industries per year. To assess how often individuals changed industries when sick, we consider a five-year window: the year before the disease onset and the four following years.[15] Off-window, individuals affected by cancer experience 0.06 different industries, while individuals with cardio-vascular disease experience 0.05 different industries. On-window, individuals with cancer worked on average in 0.07 new industries per year, while individuals with cardio-vascular diseases worked in 0.05 new industries per year. Although not reported in the table, across the nearly 10-year span, 60% of respondents changed industry at least once, and, conditional on changing, respondents changed industry at least twice. This is the variation that allows the identification of the compensation required to accept fatality risks in our panel data specification.

Table 3 reports on respondents' working conditions. One such variable is perceived physical effort at work. Almost two thirds (=60%) of respondents declared that their perceived physical effort at work was somewhat high, 8% of respondents declared it to be light and the rest declared it to be high. Another dimension captured in the Constances

the type of information included in cardio-vascular disease. The data for cardiovascular diseases includes high blood pressure, angina, myocardial infarction, stroke, arteritis of the lower limbs, other related affections, and their dates of occurrence. Cancer diseases include thyroid cancer, lung cancer, prostate cancer, colon cancer, and other cancers, as well as their dates of occurrence.

[15] As an unreported robustness check, we used a three-year window. The results are quite similar.

Table 2: Summary statistics on Constances' variables – health related

	No disease	Cancer	Cardio Vascular
Proportion of sample	87.36%	1.72%	10.92%
Number of years in panel	9.69	9.44	9.47
Age at disease onset		46.26	45.01
Number of respondents	6349	125	794
Average number of new industries per year	0.11	0.07	0.05
Off-window		0.06	0.05
On-window		0.07	0.05

Notes: Off-window/on-window is defined as a five-year window from the year before the disease onset and the four following years.

survey is satisfaction at work. Nearly 51% of respondents declared being satisfied with their current work, while the rest were not. Similarly, 49% of respondents considered that they had a high chance of obtaining a promotion in the near future.

Table 3 also reports on three type "P" variables available from CNAV. The first variable is the type of contract an individual has at each period. For a large majority of observations (88%), individuals have a full-time contract, while the remaining observations comprise temporary contracts (3%), part-time (3%) or other types of contracts (6%). The second variable corresponds to the number of quarters in a year during which an individual contributed to the public retirement fund.[16] For a quarter to be considered as valid, two conditions are required: first, a person needs to have contributed to the retirement fund; and second, to have earned at least the equivalent of 150 hours at the minimum wage (i.e., around 1,500 euros in 2015). In special cases (i.e., maternity leave, unemployment, sickness, etc.), a quarter may still be considered to be valid even if no contributions are made.[17] Finally, we only have information on individuals who contributed to the general social security scheme.[18] However, if an individual migrates between schemes, we are only able to know that they contributed to this other scheme but we are not able to observe their salary at the moment of their contribution. Fortunately, only 2% of observations belong to this category.

[16] To benefit from a retirement plan in France, individuals need to have contributed for 160 to 172 quarters, depending on their birth date.

[17] We are able to identify these cases.

[18] In France there are several health/retirement insurance schemes. The general insurance scheme referred to in this paper covers over 80% of the French population. It does not include self-employed

Table 3: Summary statistics on Constances' variables – work related

	Mean	Std. Dev.	Min	Max	Type
Physical effort at work is light	0.08	0.27	0.00	1.00	C
Physical effort at work is somewhat high	0.62	0.48	0.00	1.00	C
Physical effort at work is high	0.30	0.46	0.00	1.00	C
Satisfied with current work	0.51	0.50	0.00	1.00	C
Thinks has a high chance for a promotion	0.49	0.50	0.00	1.00	C
Full-time contracts	0.88	0.33	0.00	1.00	P
Temporary contracts	0.03	0.18	0.00	1.00	P
Part-time contract	0.03	0.18	0.00	1.00	P
Other type of contracts	0.06	0.23	0.00	1.00	P
Contributed all 4 quarters	0.80	0.40	0.00	1.00	P
Contributed 3 quarters	0.05	0.21	0.00	1.00	P
Contributed 2 quarters	0.03	0.17	0.00	1.00	P
Contributed 1 quarter	0.02	0.16	0.00	1.00	P
Contributed 0 quarter	0.10	0.29	0.00	1.00	P
All quarters in the standard scheme	0.98	0.14	0.00	1.00	P
1 quarter reported in another scheme	0.01	0.07	0.00	1.00	P
2 quarters reported in another scheme	0.00	0.05	0.00	1.00	P
3 quarters reported in another scheme	0.00	0.05	0.00	1.00	P
4 quarters reported in another scheme	0.01	0.10	0.00	1.00	P

Notes: The sample is limited to male blue-collar males aged between 20 and 59 years. Physical effort at work is light, somewhat high, or high. This variable is equal to 1 if respondents considered work to be light, somewhat high, or high, respectively, and 0 otherwise.

3.2 Industry level: CNAMTS data

Mortality and morbidity risk measures come from the French National Health Insurance Fund. The data set comprises information on around 19 million French workers covered by the National Technical Committee (NTC). This is a publicly available data set, which contains information on the total number of workers, the number of hours worked and the number of deaths on an annual basis per industry.[19]

A fatal accident is accounted for if monetary compensation is (or could have been) given to persons close to the victim. Hence, a caveat of our fatality data is that the year the death is accounted for is directly linked to the year that the compensation is allocated. Although the number of fatalities remains the same, this implies a potential mis-reporting of actual fatalities caused by work-related accidents: some deaths occurring at the end of the year could in fact be assigned to the following year.

We use information from each industry for the years 2009 to 2016 to compute fatality rates. Industry related codes are available at the four-digit level.[20] Fatal accident rates are matched to individuals' working histories in Constances, using the industry identifiers that are available in both data sets.

The fatality rate in industry j, FR_j, is computed by dividing the fatality count in a given industry, N_j, by the number of full-time equivalent employees during the calendar year, M_j, as follows:[21]

$$FR_j = \frac{N_j}{M_j} \times 100,000.$$

The fatality rates are re-expressed in terms of 100,000 FTE workers.

In principle, one could use the above equation to compute blue-collar-specific fatality rates by replacing N_j and H_j with the corresponding blue-collar-specific values (Hersh, 1998). However, while blue-collar workers' fatality numbers are provided in the CNAMTS data set, the total number of blue-collar workers by industry is not recorded at the four-digit level. Therefore, we assume that the fatality rate is proportional to the relative

or agricultural workers.

[19]The data set is available at http://www.risquesprofessionnels.ameli.fr [last visited on 9^{th} of July 2019 at 14h05].

[20]Since 2008, all statistics related to accidents follow the French National Institute of Statistics and Economic Studies (INSEE) classification: the *Nomenclature d'Activité Francaise* (NAF). NAF-related codes divide the core activity of French firms into 720 categories. Unfortunately, it is not possible to have a one-to-one match between the code used prior to 2008 and that used after 2008. As suggested by an anonymous referee, a solution may be to use a predominant match between industry classifications (NBER manufacturing database, see http://www.nber.org/nberces/). Data prior to 2009, however, is not publicly available.

[21]Information on the total number of hours worked by industry was made available starting in 2013. As a robustness check, we provide results using hours-based fatality rates. These are more tailored to the length of the exposure to risk and are thus a more accurate reflection of the worker's risk (Gentry and Viscusi, 2016). An hours-based measure would be similar to the employment-based measure if the worker group includes mostly full-time workers. The results are qualitatively similar between the two measures.

weight of blue-collar workers' deaths in the overall industry death count j. Based on this assumption, to approximate the fatality risk for blue-collar workers, we weight the average industry fatality rate, FR_j, by the share of blue-collar deaths in industry j relative to the share of blue-collar workers in industry j. Let FR_{kj} denote the fatality rate for blue-collar workers in industry j, which we compute as follows:

$$FR_{kj} = \omega_{kj} \times FR_j = \frac{\frac{N_{kj}}{N_j}}{\frac{M_{kj}}{M_j}} \times FR_j,$$

where N_{kj} denotes the number of blue-collar deaths in industry j, N_j the number of deaths in industry j, M_{kj} the number of blue-collar workers in industry j, and M_j the number of workers in industry j.

The information on the number of blue-collar workers per industry originates from INSEE. The available information, however, only contains the number of blue-collar workers for a two-digit level of aggregation (in total 38 different industries), rather than the more disaggregated four-digit level (in total 720 industries). We compute the weights, ω_{kj}, by counting employees and deaths at the two-digit level of aggregation, and assume that each ω_{kj} is equal to all industries j sharing the same two-digit level code. Our fatality risk measures are based on an eight-year average, using fatality risk from 2009 to 2016. This eight-year average is intended to smooth out irregularities in the fatality rates for cells with small employment levels, which sometimes lead to reporting zero fatalities in any given year (Kniesner et al., 2012; Gentry and Viscusi, 2016). We are able to match fatality rates to workers in specific industries using industry four-digit level identifiers.[22]

Given that our individual working history data uses information for years prior to 2009, it would be of concern to use fatality rates from later years if there were to be a sharp decreasing trend. As a validation exercise, Figure 2 illustrates the evolution of the French national fatality rates from 1955 until 2016. Fatality rates are measured by summing up the total number of fatal accidents at work in France and dividing it by the total number of workers.[23] The bold line corresponds to the yearly fatality rate, while the dotted line corresponds to the 2009-2016 average fatality rate. As Figure 2 shows, the fatality rates remain relatively constant during our period of interest.[24,25]

[22]For example, let's consider a two-digit level industry where blue-collar workers represent 30% of FTE workers. Also, let's assume that the share of blue-collar deaths in this two-digit industry is 50%. The adjustment factor is thus 0.5 / 0.3 (= 1.66), which means that 30% of the workers bear 50% of the number of work-related deaths. We then multiply the fatality rates of all four-digit industries within the two-digit industry by the 1.66 adjustment factor. Our assumption is that the distribution of blue-collar workers is the same across the four-digit industries that compose the two-digit industry.

[23]As a consequence, national fatality rates are lower than the average fatality rates taken over each industry individually.

[24]The variation across time (1955-2008) and across industry of the fatality risk is a potentially rich source of variation. However, we only have national level data for the periods prior to 2008.

[25]We provide a robustness check matching our Constances-CNAV data set with the fatality rate period (2009-2016). The results remain quantitatively and qualitatively similar.

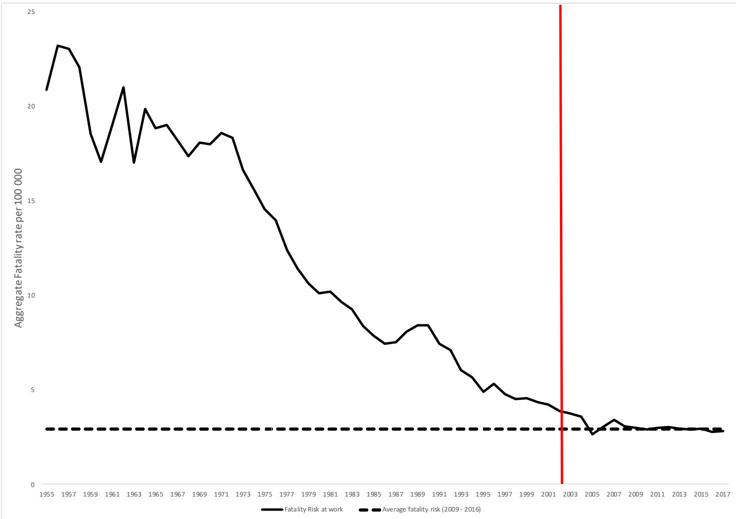


Figure 2: Evolution of the French fatality risk at work from 1955 to 2016

Table 4 reports the unadjusted (FR_j) and adjusted fatality rates (FR_{kj}) per one-digit level industry on the final Constances-CNAMTS data set. The adjusted mortality risk measures reflect the level of risk faced by blue-collar workers in each industry. Both measures are weighted averages that depend on the respondents' working history. The average unadjusted fatality risk is around 5 per 100,000 FTE workers, and there is significant variation across industries. The adjusted mortality risk measures are, unsurprisingly, higher for most industries, with an average close to 7.94 deaths per 100,000 FTE workers. The highest fatality risk is recorded in the construction sector with 9.07 deaths per 100,000 FTE workers, which reaches 12.76 after adjustment. As it involves more physical and riskier jobs, finding higher mortality risks in the construction industry is to be expected. The second highest fatality rate is in commerce, transport, accommodation, and food services, where the occurrence of fatal accidents before adjustment is 5.85 per 100,000 FTE workers, and after adjustment is 10.42 per 100,000 FTE workers. Fatality risk is 4.52 deaths per 100,000 FTE workers on average and 6.51 deaths per 100,000 FTE workers when adjusting for blue-collar workers in the scientific and technical activities.

The government, health and education industries, the real estate industry and the financial and insurance industries incur comparatively lower fatality risks, but given that most of the deaths within the industry come from blue-collar workers, the adjusted fatality risks increase markedly. On the contrary, in the information and communication industry, fatality rates decrease after adjusting for blue-collar worker exposure. The adjusted risk measure decreases, due to the comparatively small number of blue-collar worker deaths in that industry.

Table 4: Average fatality rates per industry in Constances

Industries	Fatality rate Unadjusted	Adjusted
Manufacturing	3.79	4.59
Construction	9.07	12.76
Commerce, transport, hotel & restaurant	5.85	10.42
Information and communication	2.39	0
Financial and insurance	1.92	8.52
Real estate	1.50	4.72
Scientific and technical activities	4.52	6.51
Government, health and education	1.51	7.26
Other services	1.72	2.92
Total	5.13	7.94

Notes: There are no Constances members working in the agriculture, forestry and fishing industries. Mortality rates are based per 100,000 FTE workers. Mortality rates are averaged over 2009, 2010, 2011, 2012, 2013, 2014, 2015, and 2016 risks. The analysis covers all French industries and around 19 million workers, per year. The risks presented in this table are weighted by the number of years a respondent has spent working in a given industry.

4 Empirical specification

This section describes the empirical specification used for our analysis. We first describe a hedonic wage model that only exploits cross-sectional variations. Next, we specify a hedonic wage model that exploits panel data variation.

4.1 Exploiting cross-sectional variation

The hedonic wage model adopted in this paper, in line with most of the related literature, is:

$$ln(w_i) = \alpha_0 + \alpha x_i + \beta_1 FR_{kj} + \gamma_a + \epsilon_{1i}, \tag{5}$$

where w_i corresponds to individual i's hourly gross wage; x_i is a vector of individual i's demographic characteristics; γ_a corresponds to a one-digit level industry fixed effects; FR_{kj} is the adjusted fatality rate in industry j; ϵ_{1i} rationalizes all other idiosyncratic variations. The coefficient of interest is β_1, which corresponds to the average fatality risk premium.

To assess the impact of each respondent's age and medical condition on the fatality risk premium, we use the following full-model specification:

$$ln(w_i) = \alpha_0 + \alpha x_i + \beta_z Z_i + \beta_1 FR_{kj} + \beta_2 FR_{kj} \times Z_i + \gamma_t + \gamma_a + \epsilon_{1i}, \tag{6}$$

where $Z_i = \{Mcond_i; age_{i_{30;49}}; age_{i_{50+}}\}$ is a $N \times 5$ matrix containing all N respondents' medical conditions for the two diseases considered, and three age categorical variables.[26]

[26]The excluded category corresponds to respondents below age 30.

The medical conditions variable in the cross-sectional analysis correspond to the conditions the respondent reported having at the time of the cross-sectional analysis. The interpretation of β_1 in equation (6) differs from that in equation (5). In equation (6), both coefficients correspond to the main terms and are interpreted as the monetary compensation required by a healthy (young) individual to accept a higher risk. The interaction coefficient, β_2, reports on differences with respect to the main term.

4.2 Exploiting panel variation

A caveat of the cross-sectional estimation is that it imposes a strong assumption: conditional on observed controls, x_i, the error term ϵ_{1i} is uncorrelated with the fatality rates. Although Constances proposes a large breadth of relevant control variables that might correct for a potential mis-specification, it might be the case they are still not able to fully control for relevant individual unobserved heterogeneity. There are two potential sources of individual unobserved heterogeneity that may be correlated with the fatality risks (Ashenfelter et al., 2004; Ashenfelter, 2006; Kniesner et al., 2010; Koshi et al., 2011; Kniesner et al., 2012). On the one hand, unobserved heterogeneity in personal safety productivity may be positively correlated with wages. This effect leads higher-wage workers to select what appears to be riskier jobs because the true mortality probability for the individual is lower than the measured risk. On the other hand, unobserved heterogeneity in personal safety productivity may be negatively correlated with wages. This leads higher-wage workers to select safer jobs. We relax this assumption by using panel data on workers to estimate the following model:

$$ln(w_{it}) = \alpha_0 + \alpha x_{it} + \beta_1 FR_{kj} + \gamma_t + \gamma_a + \theta_i + \epsilon_{2it}, \qquad (7)$$

where θ_i is an individual fixed effect that captures time-invariant unobservable worker-specific characteristics that affect wages and may be correlated with fatality rates; x_{it} is a vector of individual i′s time-varying demographic characteristics; ϵ_{2it} rationalizes all other idiosyncratic variations.

As our fatality risk measure is constant across years, the identification of β_1 comes from individuals switching industries that have different fatality rates. We assume, however, that conditional on individual fixed effects and on time-varying observable characteristics, the fatality rates are exogenous to ϵ_{2it}. We are not modelling individuals' joint choice on wages and industry, changing from low-wage high-risk jobs, to high-wage low-risk jobs (Abowd, McKinney and Schmutte, 2015; Lavetti, 2018; Lavetti et al., 2018). Due to our limited sample on individual-job matches, we are not able to control for this source of endogeneity.

Similarly, to assess the impact of each respondent's age and medical history on fatality risk premiums, we use the following full model specification:

$$ln(w_{it}) = \alpha_0 + \alpha x_{it} + \beta_z Z_{it} + \beta_1 FR_{kj} + \beta_2 FR_{kj} \times Z_{it} + \gamma_t + \gamma_a + \theta_i + \epsilon_{2it}, \qquad (8)$$

where $Z_{it} = \{Mhist_{it}; age_{it_{30;49}}; age_{it_{50+}}\}$ is a $N \times 5$ matrix containing all N respondents' medical history on the two diseases considered, as well as three age dummies. The identification of the β_z coefficient comes from changes in medical history and age across respondents. The identification of the β_2 coefficient comes from comparing industry changes for two types of individuals. The first type corresponds to changes of industries across time for the same individual, when this individual is sick (on-window), as compared to when the individual is not sick (off-window). The second type corresponds to changes of industries across time for the same individual who is never sick. Conditional on individual fixed effects, which capture unobservable individual heterogeneity that is fixed over time (e.g., individuals' unobserved health behavior), the off-window/on-window difference between the sick/never-sick individuals is used to capture the effect of health state on VSL. Similar to equation (6), the interpretations of β_z and β_1 depend on whether there are interactions in the model or not.

4.3 Value per statistical life estimates

The hedonic-wage equation traces out the equilibrium tangency points between firms' offer curves and workers' constant expected utility loci. That is, the VSL estimates reflect the joint influence of the supply and demand for potentially risky jobs. The VSL measures the responsiveness of wage to fatality risk.

As equation (5) uses as dependent variable the natural logarithm of wage, and since the fatality rate is measured in deaths per 100,000 FTE workers, VSL can be calculated with the following equation:

$$VSL = \hat{\beta}_1 \times \overline{wage} \times 1,607 \times 100,000, \qquad (9)$$

where the coefficient is scaled-up by 100,000 FTEs multiplied by 1,607 hours worked. All VSL estimates will be scaled down by a factor of 0.77 that corresponds to the average percentage difference between gross and net wages in France.

5 Results

Estimates of equations (5) and (6) are presented in Table 5. In the estimation, an observation is an individual. The cross-sectional sub-sample is constructed using the year when respondents were surveyed, thus using one observation per respondent.[27] This im-

[27] We could split our sample into any other year. However, we are not able to use the demographic variables reported by respondents if we use a different year than that in which the survey was

plies that the identification of the fatal risk coefficients comes only from cross-sectional variation. As respondents answered the survey in different years, all columns control for the year-of-survey. These year fixed effects are more akin to cohort fixed effects, as they only control for the year-of-survey effects (if any). Also, all specifications include one-digit level industry fixed effects.

Three models are reported in Table 5. All models include C-type demographic characteristics described in Tables 1 and 3. Model 3 examines the effects of fatality rates on wages. The coefficient on fatality rates is significantly different from zero. Thus, the implied VSL from model 3 is 6.15 million euros. Model 2 extends the analysis by including respondents' cardio-vascular disease or cancer condition. As with model 3, the fatality risk coefficient is statistically different from zero. The interpretation of the coefficient, however, differs from that of model 3. The estimated VSL for a healthy individual is 6.13 million euros. As all other coefficients are not statistically different from zero, the estimated VSL for individuals reporting either a cardio-vascular disease or a cancer disease is not statistically different from that of healthy individuals. As with previous models, model 1 includes C-type demographic characteristics, but differs by including interactions between fatality rate and respondents' age. The estimated VSL for a respondent aged below 30 is 17.67 million euros. All age-related coefficients from the interactions are statistically different from zero. The VSL estimates for individuals between 30 and 49 years and older than 50 are 3.58 million euros and 5.06 million euros, respectively.

As previously discussed, these results potentially suffer from endogeneity. Therefore, in Table 6 we examine the results exploiting respondents' work history, while allowing for individual fixed effects. Formally, this is expressed in equation (7) and equation (8). The fixed effects are aimed to capture unobserved idiosyncratic variations that may be correlated with fatal risk. Evidently, this erases a significant amount of the available variation. However, as can be seen from the standard errors, sufficient variation is left to be able to identify the effect of fatality risk on wages. The sequence of specifications in Table 6 follows the same pattern as in Table 5. Moreover, in all specifications of Table 6, we exploit all observations available for respondents. That corresponds, on average, to 10 years of observations per respondent. All models include time-varying demographic characteristics which were reported in Table 1.

Similarly to model 3, model 6 in Table 6 is the least comprehensive model, and examines the effects of fatality rates on wages. Analyzing the respondents' work history introduces more identifying variability in the data, leading to more precise estimates of the coefficients on fatality rates[28] As a result, the coefficient on fatality risk is statistically different from zero and has a smaller variance. The estimated VSL for the average respondent is 6.57 million euros. As with model 2, interactions between respondents' medical

answered.

[28]To allow for unobserved heterogeneity to be freely correlated across time for the same individual, we allow standard errors to be clustered at the individual level.

Table 5: Log wage regressions with fatality rates: cross-section

	(1)	(2)	(3)
Fatality rate	0.00957***	0.00332**	0.00333**
	(0.00285)	(0.00141)	(0.00138)
Fatality rate X age between 30 and 49	-0.00763***		
	(0.00259)		
Fatality rate X age 50 +	-0.00683**		
	(0.00305)		
Fatality rate X cancer		-0.000872	
		(0.00901)	
Fatality rate X cardio-vascular disease		0.000557	
		(0.00218)	
Has had cancer ?	-0.139	-0.131	-0.140
	(0.0853)	(0.124)	(0.0855)
Has had cardio-vascular disease ?	0.0298	0.0276	0.0321
	(0.0278)	(0.0355)	(0.0278)
Age between 30 and 49	0.275***	0.215***	0.215***
	(0.0391)	(0.0282)	(0.0283)
Age 50+	0.286***	0.233***	0.233***
	(0.0465)	(0.0330)	(0.0331)
C-type demographics	X	X	X
One-digit industry FE	X	X	X
Year-of-survey FE	X	X	X
R-squared	0.433	0.432	0.432
VSL (in millions of euros, 2015)			6.15
VSL without cancer or cardio-vascular disease		6.13	
VSL with cancer		4.52^{++}	
VSL with cardio-vascular disease		5.10^{++}	
VSL less than 30	17.67		
VSL btw 30 and 49	3.58		
VSL 50+	5.06		

Notes: VSL estimates are expressed in million euros (2015). The dependent variable is the natural logarithm of the hourly gross wage. Each model contains 4,674 observations. Industry clustered standard errors are in parenthesis. ***,**,*: significant at 1, 5 and 10% confidence levels, respectively. $^{++}$: The sum of coefficients used to compute this VSL is not statistically different from that obtained for individuals without cancer or a cardio-vascular disease.

conditions and fatality rates are introduced in model 5. The identification, however, differs from that of model 2. As we are using respondents' medical history, there is more identifying variation when using panel data than cross-section. We find a statistically significant coefficient for cardio-vascular disease, but no statistical difference for cancer. The estimated VSL for a healthy individual is around 6.48 million euros, while the estimated VSL for an individual suffering from a cardio-vascular disease is 9.46 million euros. There is no statistical difference between the VSL of a healthy individual and the VSL of an individual with cancer. As with model 1, model 4 in Table 6 documents the interactions between fatality rates and age. Following a similar pattern as in previous models, the coefficient on fatality risk is statistically different from zero. Results suggest that VSL estimates are decreasing with age. The VSL estimate for an individual below 30 is nearly 13.56 million euros. The VSL estimates for an individual between 30 and 49 years old and older than 50 are 4.38 million euros, and 4.36 million euros, respectively.

Finally, Table 7 reports on the VSL estimates from a model with a full set of interactions between individuals' medical history, their age and professional fatality risk. The results are similar to those reported for model 4 and 5, mixed together. Younger and less-healthy individuals require greater compensation than older and healthier individuals. This pattern is confirmed across the table, except for the VSL of individuals with either a cancer or a cardio-vascular disease who are older than 30 years old. The coefficient for individuals between 30 and 49 years is negative but not statistically different from zero, while we do not have any identifying variation for older individuals. This might be due to the low number of industry changes observed for individuals with either a cancer or a cardio-vascular disease in that age group.[29]

6 Robustness checks

Table 8 provides robustness checks testing different fatality risk variables and different samples, while keeping model (6) specification in Table 6. Overall, the results are robust to the different specifications and different samples. To facilitate comparability, the first column reports the same estimation results as in model (6). Column (2) reports estimates using the unadjusted count-based fatality rate from 2009 to 2016. Column (3) reports estimates from an hourly-based fatality rate using information from 2013 to 2016, while column (4) reports the unadjusted hourly-based fatality rate from 2013 to 2016. Column (5) uses the same fatality risk variable as in model (6) but restricts the sample to observations between 2009 and 2016. Column (6) uses the same fatality risk variable as in model (6) but defines blue-collar workers using the CNAV data, rather than Constances. The results are robust, and as expected from the coefficients, all models yield a similar VSL.

The next three columns report results using fatality risks aggregated to a two-digit

[29]We thank an anonymous referee for proposing this specification.

Table 6: Log wage regressions with fatality rates: panel data

	(4)	(5)	(6)
Fatality rate	0.00751***	0.00359***	0.00364***
	(0.00107)	(0.000685)	(0.000680)
Fatality rate X age between 30 and 49	-0.00502***		
	(0.000938)		
Fatality rate X age 50 +	-0.00503***		
	(0.00114)		
Fatality rate X cancer		-0.000994	
		(0.00346)	
Fatality rate X cardio-vascular disease		0.00179*	
		(0.00103)	
Has had cancer ?	-0.0775**	-0.0686*	-0.0780***
	(0.0302)	(0.0414)	(0.0303)
Has had cardio-vascular disease ?	-0.00985	-0.0232	-0.00963
	(0.0107)	(0.0145)	(0.0107)
Age between 30 and 49	0.131***	0.0937***	0.0937***
	(0.0113)	(0.00820)	(0.00820)
Age 50+	0.0396***	0.00164	0.00163
	(0.0148)	(0.0114)	(0.0114)
P-type demographics	X	X	X
Individual FE	X	X	X
One-digit industry FE	X	X	X
Year FE	X	X	X
R-squared	0.621	0.621	0.621
VSL (in millions of euros, 2015)			6.57
VSL without cancer or cardio-vascular disease		6.48	
VSL with Cancer		4.57++	
VSL with cardio-vascular disease		9.46	
VSL less than 30	13.56		
VSL btw 30 and 49	4.38		
VSL 50+	4.36		

Notes: VSL estimates are in millions of euros (2015). The dependent variable is the natural logarithm of the hourly gross wage per year. Each model contains 70,229 worker-year observations with 7,268 unique workers. Respondent clustered standards errors are in parenthesis. ***,**,*: significant at 1, 5 and 10% confidence levels, respectively. ++: The sum of coefficients used to compute this VSL is not statistically different from that obtained for individuals without cancer or a cardio-vascular disease.

Table 7: Full set of interactions between age and health states

	Age below 30	Age between 30 and 49	Age above 50	All ages
Without cancer or CVD	13.32	4.40	4.33	6.48
Cancer	36.23	11.46	2.69	4.57^{++}
CVD	31.96	8.31	7.84	9.46
Cancer and CVD	109.62	-12.94^{+}	-	-6.82^{+}
All health states	13.56	4.38	4.36	6.57

Notes: CVD stands for cardio-vascular disease. All estimates come from the same regression. The dependent variable is the natural logarithm of the hourly gross wage per year. Values are in millions of euros (2015). There are no cases in which individuals above 50: first, have overlapping episodes of cancer and cardio-vascular disease; and second, change industries. $^{+}$ The sum of coefficients used to compute this VSL is not statistically different from zero. $^{++}$: The sum of coefficients used to compute this VSL is not statistically different from that obtained for those without cancer or a cardio-vascular disease.

level industry (38 different industries). To allow for more variation, we do not control for one-digit level fixed effects. The first fatality risk variable is a blue-collar fatality risk, derived by summing the number of blue-collar workers fatalities at a two-digit level divided by the number of blue-collar workers at the same industry level. The second fatality risk variable is a fatality risk by age. As with the blue-collar fatality rate, we sum the number of workers' fatalities by age category at a two-digit level, and divide it by the number of workers in the same age category at the same industry level. Finally, the third fatality risk adjusts the second measure using the relative weight of blue-collar workers' deaths in the overall industry death count. These measures are appealing in that they ascertain the size of the fatality rates for the desired population. However, they are at a more aggregated level. As a result, the point estimate of the fatality risk using the blue-collar mortality rate is 40% larger than with the adjusted fatality rate, while the point estimate for both fatality risk per age category for younger individuals is nearly 40% lower. Lastly, the point estimates for older individuals are not statistically different than for younger individuals.

7 Conclusion

In this paper, we analyze the effects of age and baseline health on individuals' trade-offs between mortality risks and wealth. As economic theory does not provide guidance about the impact of age and baseline health on VSL, establishing the direction and magnitude of these effects remains an empirical question. This study uses a compensating wage differential approach, combining a rich population-based cohort with individual data on both work and medical history, as well as data on industry-wide fatality rates in France.

Results suggest that the average VSL estimate is close to 6.5 million euros. We find that individuals with a cardio-vascular disease consistently require greater compensation than healthy individuals for a similar exposure to mortality risk, leading to a VSL of 9.5 million euros. We find a weaker relationship for individuals with cancer. Individuals with cancer that have 50 years of age require a larger compensation than healthy individuals for a similar exposure to mortality risk, while those older than 50 years require less compensation. VSL estimates vary positively with respect to worse health status. In addition, our preferred specification also suggests that VSL estimates vary with respect to age, ranging from 13 million euros for individuals below 30 to 4 million euros for individuals above 30. These results contribute to the thin revealed preferences-based VSL literature in a European context (see Hintermann et al., 2008, for an RP in the UK context, and Baranzini and Ferro Luzzi, 2001, for an RP in a Swiss context), where the majority of the VSL literature has focused on stated preferences methods, such as Contingent Valuation Methods or Discrete Choice Experiments.[30]

Moreover, while cost-benefit analysis is increasingly used in decision-making in France, there are very few estimates of French VSL (Hammitt and Herrera-Araujo, 2016) and no studies, to the best of our knowledge, using RP approaches. As a result, today's official VSL value used by the French administration (3 million euros in 2010 euros) is derived using a benefit transfer based on an OECD meta-analysis.[31,32,33] Given the value of 6.5 million identified by using French-based trade-offs, it might be the case that French preferences for mortality risk reductions are not adequately reflected by the value currently used by the French administration. Our results also suggest that VSL varies with respect to age, and health, implying lower VSL values for older individuals and higher VSL values for sicker individuals. These findings suggest opening up the discussion, not only on the magnitude of French VSL estimates, but more generally on the legitimacy of policies aimed at differentiating VSL estimates according to age or condition.

[30] We thank an anonymous referee for proposing these references.

[31] The multi-annual French public budget programming bill of 2012 requires that all publicly funded investments should include a full socioeconomic evaluation before implementation.

[32] This value is notably higher than the previous one proposed by the Boiteux report, (1.5 million euros in 2000) or the Quinet report (3 million euros in 2010), and seems closer to the empirical values found in the CV literature (Aldy and Viscusi 2004, Robinson and Hammitt 2011). However, it is still lower than those used in some countries such as North America that primarily rely on RP estimates of the VSL. We thank an anonymous referee for pointing this out.

[33] OECD values are derived by taking the mean of a set of studies that meet a certain standard for methodological reliability. Note that all of the studies included in the OECD analysis are based on stated preferences methods.

Table 8: Robustness checks: multiple definitions of the fatality risk variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fatality rate (# workers, 2009 - 2016)	0.00364***								
	(0.000680)								
Unadjusted fatality rate (# workers, 2009 - 2016)		0.00303***							
		(0.00103)							
Fatality rate (hourly, 2013 - 2016)			0.00363***						
			(0.00068)						
Unadjusted fatality rate (hourly, 2013 - 2016)				0.00302***					
				(0.00104)					
Fatality rate (# workers, 2009 - 2016; Panel 2009-2016)					0.00362***				
					(0.000970)				
Fatality rate (# workers, 2009 - 2016; BC from CNAV)						0.00233***			
						(0.000878)			
BC fatality rate (2009 - 2016)							0.00502***		
							(0.000837)		
Fatality rate for < 30 years old (2009 - 2016)								0.00462***	
								(0.00160)	
Fatality rate for 31 to 49 years old (2009 - 2016)								-0.000816	
								(0.00149)	
Fatality rate for + 50 years old (2009 - 2016)								-0.000924	
								(0.00158)	
Unadjusted fatality rate for < 30 years old (2009 - 2016)									0.00466**
									(0.00231)
Unadjusted fatality rate for 31 to 49 years old (2009 - 2016)									0.000484
									(0.00219)
Unadjusted fatality rate for + 50 years old (2009 - 2016)									0.000487
									(0.00233)
VSL (in millions of euros, 2015)	6.57	5.47	6.55	5.47	6.54	3.70	9.06	8.34	8.41
Observations	70,229	70,229	70,229	70,229	37,050	33,578	70,229	70,229	70,229
R-squared	0.63	0.63	0.63	0.63	0.63	0.81	0.62	0.62	0.62
Number of id	7,268	7,268	7,268	7,268	6,171	9,390	7,268	7,268	7,268

Notes: BC stands for blue-collar workers. Age and blue-collar worker fatality rates are constructed by summing the number of blue-collar workers fatalities at a two-digit level, divided by the number of blue-collar workers at the same industry level. The dependent variable is the natural logarithm of the hourly gross wage per year. All models use the same specification as in model (6) from Table 6. Respondent clustered standards errors are in parenthesis. ***,** ,*: significant at 1, 5 and 10% confidence levels, respectively.

References

Abowd, J. M., McKinney, K. L. & Schmutte, I. M. (2015), 'Modeling endogenous mobility in wage determiniation', *US Census Bureau Center for Economic Studies Paper No. CES-WP-15-18* .

Alberini, A., Cropper, M., Krupnick, A. & Simon, N. B. (2004), 'Does the value of a statistical life vary with age and health status? Evidence from the US and Canada', *Journal of Environmental Economics and Management* **48**(1), 769–792.

Aldy, J. E. & Viscusi, W. K. (2004), Age variations in workers' value of statistical life, Technical report, National Bureau of Economic Research.

Aldy, J. E. & Viscusi, W. K. (2007), 'Age differences in the value of statistical life: revealed preference evidence', *Review of Environmental Economics and Policy* **1**(2), 241–260.

Aldy, J. E. & Viscusi, W. K. (2008), 'Adjusting the value of a statistical life for age and cohort effects', *The Review of Economics and Statistics* **90**(3), 573–581.

Ashenfelter, O. (2006), 'Measuring the value of a statistical life: problems and prospects', *The Economic Journal* **116**(510), C10–C23.

Ashenfelter, O. & Greenstone, M. (2004), 'Estimating the value of a statistical life: The importance of omitted variables and publication bias', *American Economic Review* **94**(2), 454–460.

Baranzini, A., Luzzi, G. F. et al. (2001), 'The economic value of risks to life: evidence from the swiss labour market', *Revue Suisse d Economie Politique et de Statistique* **137**(2), 149–170.

Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Pearce, D. W. et al. (2002), 'Economic valuation with stated preference techniques: A manual.', *Economic valuation with stated preference techniques: a manual.* .

Bleichrodt, H. & Eeckhoudt, L. (2006), 'Willingness to pay for reductions in health risks when probabilities are distorted', *Health Economics* **15**(2), 211–214.

Cameron, T. A. & DeShazo, J. (2013), 'Demand for health risk reductions', *Journal of Environmental Economics and Management* **65**(1), 87–109.

Cropper, M., Hammitt, J. K. & Robinson, L. A. (2011), 'Valuing mortality risk reductions: progress and challenges', *Annu. Rev. Resour. Econ.* **3**(1), 313–336.

Davis, L. W. (2004), 'The effect of health risk on housing values: Evidence from a cancer cluster', *American Economic Review* **94**(5), 1693–1704.

Dorman, P. & Hagstrom, P. (1998), 'Wage compensation for dangerous work revisited', *ILR Review* **52**(1), 116–135.

Ehrlich, I. & Yin, Y. (2005), 'Explaining diversities in age-specific life expectancies and values of life saving: a numerical analysis', *Journal of Risk and Uncertainty* **31**(2), 129–162.

Evans, M. F. & Schaur, G. (2010), 'A quantile estimation approach to identify income and age variation in the value of a statistical life', *Journal of Environmental Economics and Management* **59**(3), 260–270.

Fairris, D. (1989), 'Compensating wage differentials in the union and nonunion sectors', *Industrial Relations: A Journal of Economy and Society* **28**(3), 356–372.

Finkelstein, A., Luttmer, E. F. & Notowidigdo, M. J. (2013), 'What good is wealth without health? the effect of health on the marginal utility of consumption', *Journal of the European Economic Association* **11**(suppl_1), 221–258.

Gayer, T., Hamilton, J. T. & Viscusi, W. K. (2002), 'The market value of reducing cancer risk: Hedonic housing prices with changing information', *Southern economic journal* pp. 266–289.

Gentry, E. P. & Viscusi, W. K. (2016), 'The fatality and morbidity components of the value of statistical life', *Journal of health economics* **46**, 90–99.

Hammitt, J. K. (2002), 'Qalys versus wtp', *Risk Analysis* **22**(5), 985–1001.

Hammitt, J. K. & Herrera-Araujo, D. (2018), 'Peeling back the onion: Using latent class analysis to uncover heterogeneous responses to stated preference surveys', *Journal of Environmental Economics and Management* **87**, 165–189.

Hammitt, J. K. & Robinson, L. A. (2011), 'The income elasticity of the value per statistical life: transferring estimates between high and low income populations', *Journal of Benefit-Cost Analysis* **2**(1), 1–29.

Hersch, J. (1998), 'Compensating differentials for gender-specific job injury risks', *The American Economic Review* **88**(3), 598–607.

Hintermann, B., Alberini, A. & Markandya, A. (2010), 'Estimating the value of safety with labour market data: Are the results trustworthy?', *Applied Economics* **42**(9), 1085–1100.

Johansson, P.-O. (2002), 'On the definition and age-dependency of the value of a statistical life', *Journal of Risk and Uncertainty* **25**(3), 251–263.

Jones-Lee, M. W. (1989), *The economics of safety and physical risk*, Basil Blackwell.

Kniesner, T. J., Viscusi, W. K., Woock, C. & Ziliak, J. P. (2012), 'The value of a statistical life: Evidence from panel data', *Review of Economics and Statistics* **94**(1), 74–87.

Kniesner, T. J., Viscusi, W. K. & Ziliak, J. P. (2006), 'Life-cycle consumption and the age-adjusted value of life', *The BE Journal of Economic Analysis & Policy* **5**(1).

Kniesner, T. J., Viscusi, W. K. & Ziliak, J. P. (2010), 'Policy relevant heterogeneity in the value of statistical life: New evidence from panel data quantile regressions', *Journal of Risk and Uncertainty* **40**(1), 15–31.

Kochi, I. & Taylor, L. O. (2011), 'Risk heterogeneity and the value of reducing fatal risks: further market-based evidence', *Journal of Benefit-Cost Analysis* **2**(3), 1–28.

Lavetti, K. (2018), 'The estimation of compensating wage differentials: Lessons from the deadliest catch', *Journal of Business & Economic Statistics* (just-accepted), 1–36.

Lavetti, K. & Schmutte, I. (2018), 'Estimating compensating wage differentials with endogenous job mobility', *technical report http://www.kurtlavetti.com/CDEMvc.pdf* (retrieved July 17 2019).

Lee, J. & Taylor, L. O. (forthcoming), 'Randomized safety inspections and risk exposure on the job: quasi-experimental estimates of the value of a statistical life', *American Economic Journal Economic Policy* .

O'Brien, J. (2018), 'Age, autos, and the value of a statistical life', *Journal of Risk and Uncertainty* **57**(1), 51–79.

Rheinberger, C. M., Herrera-Araujo, D. & Hammitt, J. K. (2016), 'The value of disease prevention vs treatment', *Journal of health economics* **50**, 247–255.

Sloan, F. A., Viscusi, W. K., Chesson, H. W., Conover, C. J. & Whetten-Goldstein, K. (1998), 'Alternative approaches to valuing intangible health losses: the evidence for multiple sclerosis1', *Journal of health Economics* **17**(4), 475–497.

Smith, V. K., Evans, M. F., Kim, H. & Taylor Jr, D. H. (2004), 'Do the near-elderly value mortality risks differently?', *Review of Economics and Statistics* **86**(1), 423–429.

Treasury, H. M. (2003), *The green book: appraisal and evaluation in central government*, London TSO.

Viscusi, W. K. (2004), 'The value of life: estimates with risks by occupation and industry', *Economic Inquiry* **42**(1), 29–48.

Viscusi, W. K. & Aldy, J. E. (2003), 'The value of a statistical life: a critical review of market estimates throughout the world', *Journal of risk and uncertainty* **27**(1), 5–76.

Viscusi, W. K. & Aldy, J. E. (2007), 'Labor market estimates of the senior discount for the value of statistical life', *Journal of Environmental Economics and Management* **53**(3), 377–392.

Viscusi, W. K. & Evans, W. N. (1990), 'Utility functions that depend on health status: estimates and economic implications', *The American Economic Review* pp. 353–374.