WORKING PAPERS
Documents de travail i3

Are older nuclear reactors less safe? Evidence from incident data in the French fleet

Romain Bizet
CERNA – Centre for industrial economics, i3 UMR CNRS
MINES ParisTech, PSL Research University
romain.bizet@mines-paristech.fr

Petyo Bonev
CERNA – Centre for industrial economics, i3 UMR CNRS
MINES ParisTech, PSL Research University
petyo.bonev@mines-paristech.fr

François Lévêque
CERNA – Centre for industrial economics, i3 UMR CNRS
MINES ParisTech, PSL Research University
francois.leveque@mines-paristech.fr

Working Paper 17-CER-02
Février, 2017

Pour citer ce papier / How to cite this paper : Bizet, R. ; Bonev, P. & Lévêque, F. (2017) Are older reactors less safe? Evidence from incident data in the French nuclear fleet. i3 Working Papers Series, 17-CER-02.

L'institut interdisciplinaire de l'innovation

(UMR 9217) a été créé en 2012. Il rassemble :

- les équipes de recherche de MINES ParisTech en économie (**CERNA**), gestion (**CGS**) et sociologie (**CSI**),
- celles du Département Sciences Economiques et Sociales (**DSES**) de Télécom ParisTech,
- ainsi que le Centre de recherche en gestion (**CRG**) de l'École polytechnique,

soit plus de 200 personnes dont une soixantaine d'enseignants chercheurs permanents.

L'institut développe une recherche de haut niveau conciliant excellence académique et pertinence pour les utilisateurs de recherche. Par ses activités de recherche et de formation, i3 participe à relever les grands défis de l'heure : la diffusion des technologies de l'information, la santé, l'innovation, l'énergie et le développement durable. Ces activités s'organisent autour de quatre axes :

- Transformations de l'entreprise innovante
- Théories et modèles de la conception
- Régulations de l'innovation
- Usages, participation et démocratisation de l'innovation

Pour plus d'information : http://www.i-3.fr/

The Interdisciplinary Institute of Innovation

(UMR 9217) was founded in 2012. It brings together:

- the MINES ParisTech economics, management and sociology research teams (from the CERNA, CGS and CSI),
- those of the Department of Economics and Social Science (DSES) at Télécom ParisTech,
- and the Management Research Center (CRG) at Ecole Polytechnique,

meaning more than 200 people, including 60 permanent academic researchers.

i3 develops a high-level research, combining academic excellence and relevance for the end users of research. Through its teaching and research activities, i3 takes an active part in addressing the main current challenges: the diffusion of communication technologies, health, innovation, energy and sustainable development. These activities are organized around four main topics:

- Transformations of innovating firms
- Theories and models of design
- Regulations of innovation
- Uses, participation and democratization of innovation

For more information: http://www.i-3.fr/

Are Older Nuclear Reactors Less Safe? Evidence from Incident Reports in the French Fleet*

Romain Bizet[†1], Petyo Bonev[1], and François Lévêque[1]

[1]Mines ParisTech, Centre for Industrial Organization, PSL Research University, CNRS UMR 9217

February 10, 2017

Abstract

This paper studies the impact of age and reactor technology on safety in the French nuclear fleet between 1997 and 2015. We use a novel dataset encompassing over 19,000 nuclear safety events declared by plant managers to the French regulatory agency. A major problem for evaluating the effects of age and technology is that declarations of safety events are influenced by the plant managers' level of transparency. We deal with this problem by restricting the analysis to the occurrences of particular types of events, such as automatic shut-downs. These events, due to their technical specifics, exhibit perfect detection and declaration rates. We obtain the following results. First, technology has a strong impact on reactor safety. Second, age has a significant and technology-specific effect on reactor safety. For instance, while in 1997 one additional year of age led to a 15% increase in the expected number of automatic shut-downs among the 900 MW reactors, this number was reduced to only 6% in 2014. In comparison, the 1450 MW reactors undergo a significantly larger number of automatic shut-downs, but age has a smaller effect on their rate of occurrence. Finally, we find that local transparency, defined as both detection abilities, reporting behaviours and declaration guidelines, plays a significant role in the explanation of the observed variations of declarations of events.

Keywords: nuclear power, safety, age, technology, reactor design, incident data.

1 Introduction

In this paper, we empirically evaluate the effect of age and technology of nuclear reactors on their safety, using data on significant safety incidents reported between 1997 and 2015 in the 58 French nuclear reactors. From a policy-oriented perspective, this question comes at a timely moment, as the oldest French power stations are reaching their fortieth anniversary - their initial maximum lifespan - and are now seeking for licence extensions, which led EDF (France's single nuclear utility), safety regulators and policy-makers to argue over the effect of age on the safety of these plants. An important question of the policy debate surrounding

*We thank the audience of the Young Energy Economist and Engineers seminar held in Edinburgh University for their questions and remarks regarding this paper.

†Contact details: Mines ParisTech, Centre for Industrial Economics, 60 Boulevard Saint-Michel, 75006, Paris. romain.bizet@mines-paristech.fr

nuclear power is thus whether nuclear power stations ought to be shut-down on the sole basis of their age. Despite this high political and social importance, this question has not been answered in a satisfactory way by the existing literature. The available data on nuclear safety is often of limited amount or quality, since major safety accidents are rare. Researchers have therefore mostly relied on extended datasets that included accidents from both nuclear reactors and fuel cycle facilities (Hirschberg et al., 2004; Sovacool, 2008; Hofert and Wüthrich, 2011; Wheatley et al., 2016a,b). Due to the scarcity of these accidents, these datasets contained at most 216 observations, which limits the possibilities of inferring statistically significant results regarding the safety of various nuclear technologies, or to study its variations with the age of nuclear reactors.

The statistical estimation of the variations of nuclear safety across reactors and years faces a methodological challenge. In industries characterized by rare occurrences of catastrophic accidents, observed safety indicators may be prone to measurement errors. This is particularly true when safety indicators are based on incident information reported by the industry, which is the case of the nuclear industry.[1] If there is a positive probability that these events remain undetected, or if firms have incentives not to report them[2], then a spurious effect in the data could bias the estimates, due to the unobservability of both detection abilities and compliance with declaration criteria. Indeed, the ability to detect an event might vary with the design of nuclear reactors, while the propensity of firms to comply with declaration criteria might vary over time with the evolution of the stringency of the safety standards. An instance of this spurious effect was described by Rose (1990) within the airline industry, in which safety incidents are also used as a measure of safety. An additional bias precluding the identification of changes in the observed frequencies of events as a measure of safety development stems from the fact that the criteria leading a firm to report an incident might vary over time. In the following of this paper, we refer to both potential measurement errors (due to non-declaration or non-detection of events) and changes in regulatory declaration criteria as *transparency issues*.

Our paper contributes to the empirical literature on the effect of age and technology on nuclear safety in several ways. First, we introduce a novel dataset which contains all nuclear safety incidents reported between 1973 and 2015 in French nuclear reactors by France's single

[1]As an example, the World Association of Nuclear Operators defines a set of key safety performance indicators, in which are included the annual numbers of automatic and manual scrams (shut-down of a nuclear reactor).

[2]The fear of more stringent regulatory standards or of public backlashes embodies these incentives. A canonical example of this situation in the nuclear industry is the occupation of the French Fessenheim power plant in 2016 by Greenpeace after a German newspaper claimed that an incident had been dramatically understated by the French nuclear safety authority in 2014.

utility EDF. This dataset gathers detailed information on more than 19,000 significant safety incidents. These events represent the most significant departures from the general rules of operation of a nuclear power reactor, and correspond to a set of compulsory declaration criteria defined by the French Nuclear Safety Authority (ASN). They are analysed on a case-by-case basis by both plant managers from EDF, and by experts from IRSN,[3] the technical body of ASN. These case-studies aim to identify organizational or technical weaknesses in nuclear stations, and to foster the exchange of best practices across the nuclear fleet. Yet, to the best of our knowledge, no statistical analysis of these events has ever been conducted. We conduct a comprehensive descriptive analysis of this data and several interesting patterns emerge. In particular, different types of events appear to be characterized by various time-trends, while the annual counts of reports of events per reactor seem to be unevenly distributed across reactors of different age and technology.

Second, we use this dataset to estimate the causal effect of the age of a reactor and of its technical properties on its safety level, measured as the number of safety events for a fixed period of time. In order to preclude the aforementioned possible measurement error, we identify a subset of events that exhibits perfect detection and declaration rate. They owe this property either to their impact on the electricity output (e.g. automatic shut-downs) or to extraordinary auditing measures exerted by the regulator (e.g. events that require the unplanned use of the safeguard mechanisms of the power station). This strategy is related to the strategy used in Hausman (2014), who uses automatic shut-downs, among other safety metrics, to evaluate the effect of economic incentives on safety in the US nuclear market. Furthermore, in order to preclude spurious causal effect due to changing regulation, we restrict the period of observation to 1997 - 2015. During this period, the criteria leading to the declaration of these events remained unchanged. By considering events characterized by perfect detection and declaration, and on a period with constant declaration criteria, all variations in the number of declared events must be due to variations in reactor safety. This is a novel approach for the analysis of nuclear safety. In addition to estimating the effect of age and technology, we are able to study the evolution of safety events over calendar time. Such an evolution is likely to reflect technological progress over time. It thus provides further insights on the importance of technology for nuclear safety.

Third, we analyse whether transparency as defined above affects the declaration of significant safety events. To do so, we compare the results obtained on the set of perfectly detected and declared events with the results of a similar regression run on the whole set of nuclear

[3] IRSN stands for Institute for Radio-protection and Nuclear Safety.

incidents, arguing that if transparency has no effect on the declaration process, the results obtained on both sets ought to be similar. This approach is related to the approach in Rose (1990) who analyses the effect of market deregulation on airline safety in the eighties.

Our results first indicate that technology has a strong impact on reactor safety. In particular, reactors of more recent technological design declare significantly higher numbers of perfectly detected and declared events each year. We also find that age has a negative effect on the reactor's safety, but that this effect decreases over time and is heterogeneous across the different designs of reactors. For instance, for France's 900 MW reactors, one additional year of age led to a 15% increase in the expected number of automatic shut-downs per year in 1997, while this figure fell to only 6% in 2014. For France's 1450 MW reactors, the effect of age changed sign over this same period of time: in 1997, age had no significant effect on the total number of automatic shut-downs, while their frequency decrease by 8% per year of age in 2014. In other words, our results suggest that although the ageing of reactors can be slightly detrimental to their safety levels, significant progresses in the management of ageing have been achieved. Finally, it appears that transparency issues bias the estimation of the effect of age or technology on safety, by reducing the heterogeneities of declaration observed on the restricted set of perfectly detected and declared event.

The paper is organized as follows. Section 2 describes the French declaration process and conducts a descriptive analysis of our dataset. Section 3 presents the identification strategy and empirical specifications. Section 4 exposes our results and section 5 concludes.

2 Significant safety events in the French fleet

2.1 Institutional set-up and data

The French nuclear fleet is constituted of 58 pressurized water reactors (PWR), located in 19 power stations, and owned by a single utility (EDF). These reactors were built in separate phases from the late 1970s to the late 1990s. Several groups of reactors can be distinguished, as their technological features evolved between each construction phases. For instance, reactors differ in their nominal capacity, the nature of their fuel, or in their ability to perform load-following. In this paper, we distinguish two particular groups of nuclear reactors: capacity groups, and conception groups. The fleet can be split within three capacity groups, each of which being constituted of one or several conception groups, as summarized in table 1.

Most importantly, the 900 MW[4] group contains reactors of three different conceptions (CP0, CP1 and CP2 reactors)[5]. Likewise, the 1300 MW group contains reactors of two different conceptions (P4 and P'4 reactors).

Table 1: The French fleet, by conception and nominal capacity levels.

Capacity	Conception	Power stations	Reactors	Construction
900 MW	CP0	2	6	1971-1979
	CP1	4	18	1974-1985
	CP2	3	10	1976-1988
1300 MW	P4	3	8	1977-1986
	P'4	5	12	1980-1992
1450 MW	N4	2	4	1984-2000

note: Construction phases span from the beginning of the construction of the first reactor and until the connection of the last one.

The safety of French nuclear power stations is regulated by the Nuclear Safety Authority (ASN in the following), which sets regulatory standards for the nuclear operator, EDF. In particular, the regulatory standards include the mandatory declarations of a list of particular situations, or events, that depart from the general rules of operations of nuclear power stations. The ASN enforces these standards through periodic inspections. Every ten year, each reactor is thoroughly reviewed, and mandatory safety investments are defined by the ASN in order to pursue operation. These investments consist, for instance, in the replacement of materials which are assumed to have been deteriorated in a way that precludes their further safe use. In addition to these long-term visits, planned and unplanned routine inspections are conducted several times a year. In particular, these visits are used by the inspectors to verify that plant managers do comply with the declaration guidelines set by the safety authority.

The aim of these declarations is to share experience and best practices among reactors, and to detect generic defaults in the design of the reactors. Events that ought to be declared are defined according to a set of 10 criteria established by the regulator. Once an event is detected in a reactor, plant managers have two weeks to provide the regulator with a detailed a summary of the event and an analysis of its causes and consequences. During the routine inspections of a reactor, inspectors can go through the reactor's history of situations which were not deemed significant enough for declaration. If inspectors disagree with the plant managers' judgement, they can force them to declare the event.[6] The definitions of the declaration

[4]MW stands for MegaWatt. A 900 MW reactor is thus a reactor whose nominal capacity of production is 900 MegaWatts.

[5]Each conception of reactors is characterized by specific technological features.

[6]Not declaring an event which falls under the scope of the 10 regulatory declaration criteria is a violation that would allow the regulator to engage in repressive actions against the plant manager. In practice, the regulator seldom uses his right to litigate EDF.

criteria evolved over time, leading to variations in the annual number of declarations as well as in the descriptive quality of each declaration.

Our dataset contains all significant safety events declared from 1973 to 2015 in the French nuclear fleet. It includes over 19.000 events, and was provided to us by the French nuclear safety regulator.[7] Each event is characterized by a set of variables, which contain information on the location and date of declaration of each event, the nature of the components, materials and systems of the reactor affected by the events, their level on the International Nuclear Event Scale[8] (INES), the declaration criteria associated with the events, the state of the reactor at the time of detection (such as production, refuelling or maintenance) and several elements regarding the causes and consequences of each event.

Not all of these events are used for our analysis. First, generic events are excluded.[9] We also discarded the events declared during the construction phases of each reactor, as our interest in this paper is the safety of operated reactors. Another issue was raised by "common" events, defined as events that impact systems used by multiple reactors within a given power station. When aggregating safety events into total numbers of declarations per reactor and per year, we chose to duplicate these common events in order to associate them with each affected reactor. Once these modifications have been implemented in the dataset, we obtain 21,628 events over the period 1978-2015.

2.2 Descriptive statistics

To provide a short description of this dataset, we first construct time-series describing the evolution of the declarations of particular categories of events over time. Figure 1 presents the evolution of four categories of events from 1977 to 2015: the yearly number of events declared in the fleet, INES-1 events[10], automatic reactor shut-downs, and events that require the use of the safeguard mechanisms of a nuclear reactor.

The upper left graph on figure 1 shows that the total number of declarations per year is increasing over time. This might be due to the increasing number of nuclear power stations in operation between 1977 and 1999, and to the evolution of the declaration guidelines provided by the regulator, which became more and more stringent over time. Another interesting

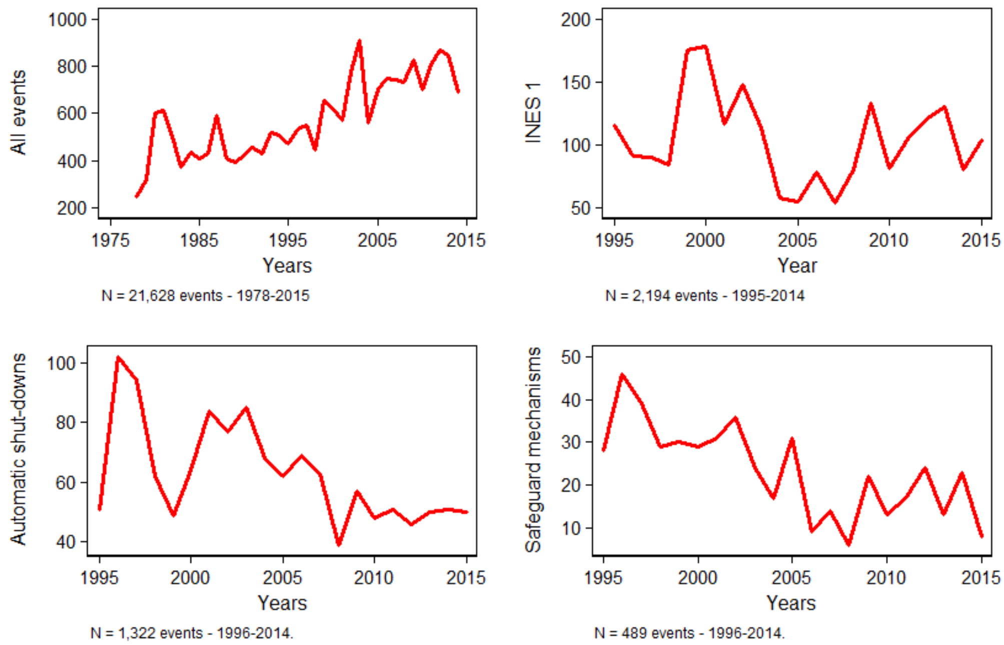
[7]Events declared during the lifetime of permanently shut-down reactors are not included in our dataset.

[8]The International Nuclear Event Scale is a severity indicator for nuclear events, defined by the International Atomic Energy Agency.

[9]Generic events consist in detections of conception failures which are specific to either a group of reactors, or to the whole fleet. According to the regulator, these events capture specific efforts made by EDF to increase his knowledge of the conception of his reactors, as well as their reliability.

[10]An INES-1 event is here defined as an event whose level on the INES scale is equal to 1. These are part of the severe events contained in our data, as most of the events in our dataset are ranked 0 on the INES scale.

Figure 1: Annual declarations of four sub-categories of events in the French fleet



takeaway from these first descriptive statistics is the observation that different categories of events can be subject to different time-trends. While the number of automatic shut-downs or events requiring the use of safeguard mechanisms have been steadily decreasing over time, INES-1 events exhibit less clear patterns. These differences might suggest that plant managers may be able to exert specific efforts to limit the occurrences of particular types of events. For instance, automatic shut-downs stop the production of electricity, creating a direct monetary incentive for EDF to reduce their frequency.

Next, we compare the frequencies of safety events across reactors or groups of reactors. To do so, we first construct statistics describing reactor.years. A reactor.year summarizes all the events occurring in a given reactor during a given year: for instance, the Gravelines-5-2004 reactor.year is constituted of all the events declared in the fifth reactor of the Gravelines power station in 2004. Figure 2 presents box plots showing how the annual declarations of events are distributed across four definitions of groups of reactors. In each box plot, a reference group is indicated in green. Red whisker boxes identify groups of reactors that present a significantly higher annual average number of events with respect to this reference group, at the 5% level of confidence.

On the first line of figure 2, it appears that the reactors that belong to the most recent capacity group (1450 MW), declare a significantly larger average number of events when compared to the 900 MW. Some heterogeneity appears within capacity groups when split into the six underlying conception groups. Within the 900 MW group, the $CP0$ design declares a significantly higher number of events than the other two designs. The same conclusion can be drawn for the reactors of the $P4$ conception with respect to their younger siblings of the $P'4$ conception.

The bottom-left graph on figure 2 is dedicated to France's regulatory subdivisions.[11] It appears that only the reactors overseen by the Bordeaux-based regulator (e.g. the 8 reactors located in the Blayais, Golfech and Civaux power stations) exhibit a significantly lower number of annual declarations than the power stations overseen by the Lille-based regulator. Reactors from the other five territorial subdivisions of the ASN do not significantly differ in declarations from the Lille-reference.

On the bottom-right graph, four age groups are defined by the upper decade of operation of a reactor at a given year. For instance, the Fessenheim-2-2002 reactor.year, during which the Fessenheim-2 reactor was aged 25, is associated with the third decade of operation of this reactor. Reactors in their third and fourth decade of operation declare significantly larger numbers of events than reactors in their first decade of operation. It is interesting to notice the relative under-dispersion of the fourth-decade group with respect to the three younger groups. This could be due to a relative lack of observation of fourth-decade reactor.years. Indeed, the current mean-age of the French nuclear fleet is 30 years, and as of 2014, the first, second and third decade groups respectively count 743, 563 and 414 observed reactor.years, whereas only 45 reactor.years of the fourth decade have been observed.

3 The impact of age and technology on nuclear safety

3.1 Identification strategy

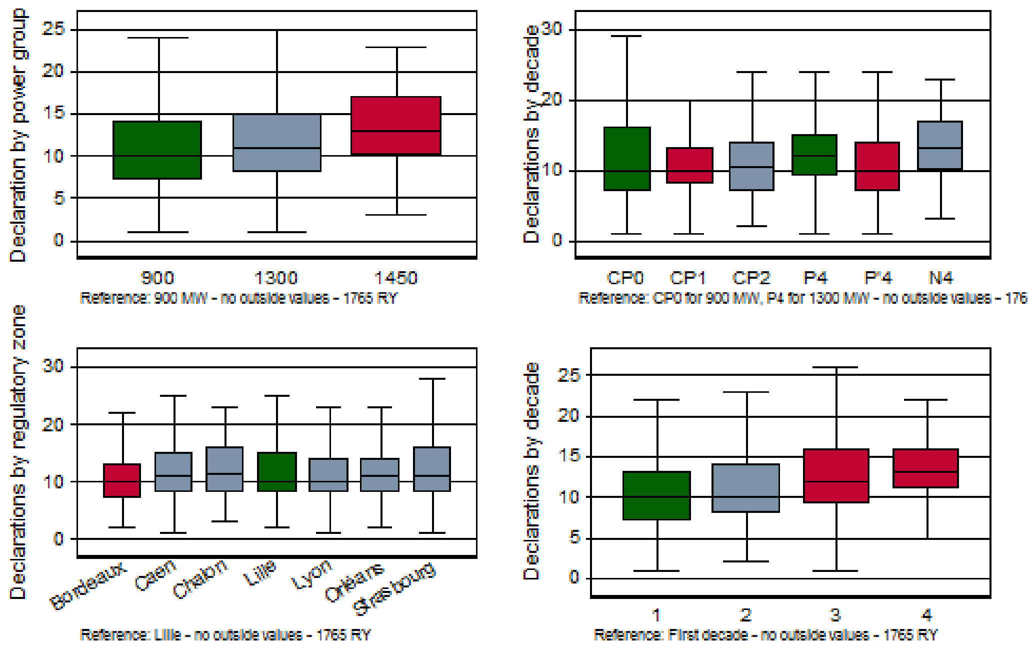
Assume Y is a binary random variable such that $Y = 1$ when a safety event occurs. Our objective is to measure the evolutions of *safety*, which we define here as $\mathbb{P}\{Y = 1\}$. The rationale for this definition is twofold. First, it is consistent with the definition of safety as a probability of occurrence of events which may harm people or goods, which has become usual in the economics literature.[12] In addition, reducing the likelihood of these incidents has a direct impact on the likelihood of major accidents, as these accidents are often composed of a combination of individually minor events. Consider the model

$$\mathbb{P}\{Y = 1|W, C\} = \mathbb{E}\{Y|W, C\} = g(W, C), \tag{1}$$

[11] The ASN delegates the duty of inspection to its 7 territorial subdivisions, who have some level of discretion in their interaction with the plant managers.

[12] See for instance Shavell (1984); Hansson and Skogh (1987); Faure and Skogh (1992) or Laffont (1995).

Figure 2: Annual declarations by reactor for different groups of reactors



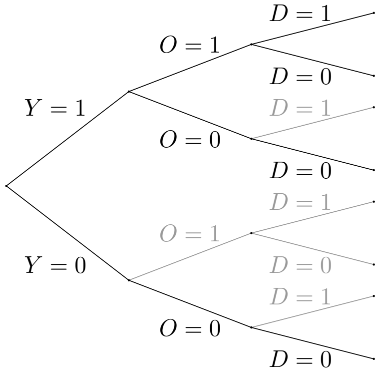
where the random vector W contains observed factors of safety such as age and technology, and the random variable C is unobserved. Examples for factors in C are the unobserved (reactor-specific) ability to detect an event, and the propensity with wich an operator would be willing to declare a detected event. g is the regression function in the population. We are interested in the marginal effect of a covariate W_j (e.g. technology) on the likelihood that an event occurs: $\frac{\partial \mathbb{E}\{Y|W,C\}}{\partial W_j}$. If, however, C is related to W, then the observed regression function $\mathbb{E}\{Y|W\}$ will capture this effect, which will lead to a bias, $\mathbb{E}\{Y|W,C\} - \mathbb{E}\{Y|W,C\} \neq 0$. To formalize this, assume O is a binary random variable such that $O = 1$ if a safety event is observed, or detected, by the plant manager. $\mathbb{P}\{O = 1|Y = 1\}$ represents the plant manager's *ability to detect* events. Assume D is a binary random variable such that $D = 1$ if a safety event is declared by the plant manager to the regulator. D captures the plant manager's *propensity to declare* events. Given the description of the declaration process and its regulatory oversight, we assume that type I errors cannot occur:

$$\mathbb{P}\{O = 1|Y = 0\} = \mathbb{P}\{D = 1|O = 0\} = 0. \tag{A1}$$

Intuitively, assumption (A1) means that the plant manager cannot detect an event which did not occur, nor can he declare an event which has not been previously detected. The event tree presented on figure 3 sums up the possible states of nature.

With this notation, the only data observable to the econometrician are the reports, which

Figure 3: Event tree of the generating process of reported events



occur only when the realized state is $\{Y = 1, O = 1, D = 1\}$. Therefore, transparency can bias our study through two channels. First, events can fail to be detected. Second, detected events can fail to be declared. Thus, provided constant declarations criteria, a change in the observed number of reports can be explained by three factors: a true change in safety, a change in the detection ability of the plant managers, or a change in their level of compliance with declaration guidelines (propensity to declare). If these channels are related to technology and/or age, then the estimated effects of these variables will be biased. In addition to these issues, changes in the regulatory framework over time might also affect the rate of declaration of events, namely through the definition of an event (and hence through the requirement on what should be declared). When all observations are pooled together and such regulatory changes ignored, different reactors could reveal different observed frequencies of events simply because the measurements were taken at different points in time. Thus, because calendar time is related to age, ignoring changes in the regulatory framework over time could lead to a spurious effect of age in the data.

Our identification strategy consists in finding a subset of the sample, for which the aforementioned problems are precluded. Assume first that the definitions of significant safety events are constant over time. We define an event to be *perfectly detectable and declared (PDD)* if the following assumption is fulfilled:

$$\mathbb{P}\{O = 1|Y = 1\} = \mathbb{P}\{D = 1|O = 1\} = 1. \tag{A2}$$

We denote the subset of all PDD events as Θ_{PDD}. We obtain the following result:

Lemma 3.1. *Assume that the regulatory standards are fixed and that assumptions* (A1) *and* (A2) *hold. Then it holds*

$$\mathbb{E}\{Y|W, C\} = \mathbb{E}\{Y|W\}. \tag{2}$$

The proof can be found in the appendix.

Thus, according to lemma (3.1), the model $g(W, C)$ can be uncovered from the observed data $\mathbb{E}\{Y|W\}$ for the subset of PDD events. On this subset, variations in the frequencies of occurrence of safety events are necessarily caused by variations in safety. We describe the set Θ_{PDD} In the following section.

To check the external validity of this strategy, we also measure the variations of $\mathbb{E}\{Y|W\}$ on the unrestricted set of safety events declared between 1997 and 2015. A discussion of this test of external validity is provided at the end of section 4.2.

3.2 Perfectly detected and declared events

For any particular type of safety incidents, there are two mutually independent conditions which are sufficient to guaranty perfect detection and declaration. First, events that have a direct effect on the electrical output of a power station cannot be undetected or hidden as the Transportation System Operator[13] monitors the electric production of each power station. Second, events which are subject to particular auditing efforts from the inspectors ought to be declared truthfully, when the probability of uncovering non-compliant behaviours is very high, which would reduce the incentives for plant managers to deviate.

Jointly with experts from the French safety regulator, we identified two categories of events that satisfy at least one of these two conditions. First, automatic shut-downs of reactors have an impact on the electrical output of the power station. Using automatic shut-downs as a proxy for nuclear safety is also supported by the fact that the annual number of automatic shut-downs is retained by the World Association of Nuclear Operators as one of their safety performance indicators[14]. They have also been used by Hausman (2014) as a proxy for nuclear safety. Second, events that require the unplanned use of the safeguard mechanisms of the power station (safeguard events in the following) are subject to specific auditing efforts during the inspection of the power stations. The relative severity of these events makes them rather easy and natural targets for the ASN inspectors during the routine inspections of power stations.[15]

Both automatic shut-downs and safeguard events are being systematically tagged in the dataset since 1995. In order to preclude recording errors, we exclude the first two years of observations (1995 and 1996). We thus restrict our analysis to the period 1997-2015. A

[13]In France, until 2000, the electricity transportation network was managed by EDF. Since 2000, transmission and production have been unbundled, and the transmission network has been handled by a single operator (RTE), which remains a subsidiary of EDF.

[14]See for instance WANO's yearly performance reports on their website.

[15]Interviews conducted with both the ASN and EDF seem to suggest that making the assumption that safeguard events are perfectly declared is reasonable.

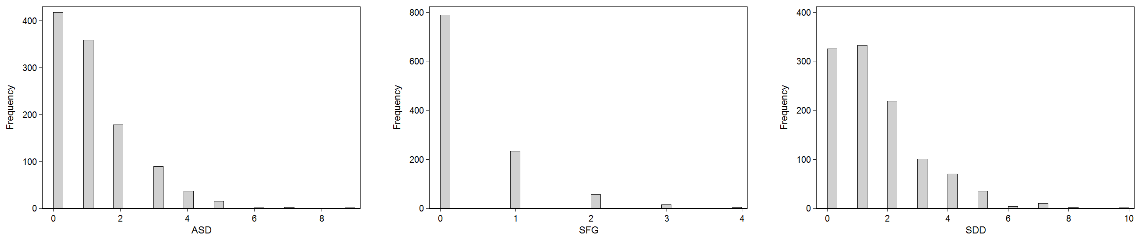
necessary condition for the validity of our strategy is the time-invariance of the definition of the criteria over the time-period considered. As the conditions under which a reactor is supposed to be automatically shut-down have remained constant over the past twenty-years, and by definition of safeguard events, it appears that the condition is met for automatic shut-downs and safeguard events. [16] Table 2 and figure 4 provide some descriptive statistics associated with the annual counts of automatic shut-downs and safeguard events.

Table 2: Summary statistics for our four dependent variables

Variable	Definition	Mean	Std. Dev.
ASD	Automatic shut-downs declared during year t in reactor r	1.122	1.233
SFG	Events requiring the use of safeguard mechanisms declared during year t in reactor r	0.377	0.693
PDD	Perfectly detected and declared events (ASD + SFG) declared during year t in reactor r	1.499	1.509
ALL	All events declared during year t in reactor r	12.256	5.105

1,100 reactor.years - 1997-2015.

Figure 4: Occurrences of perfectly detected and declared events per reactor.year



3.3 Empirical specifications

Let Y_{it} denote the counts of safety events declared during year t in reactor i, $t \in \{1997, \ldots, 2015\}$, $i = 1, \ldots, 58$. Further, let AGE_{it} be the age of reactor i in year of observation t, and let X_{it} denote a set of reactor and year specific control variables, such as whether the reactor is a first-of-a-kind or a first-of-a-site, or the size of the plant in which a reactor is located, measured by the number of reactors in the plant.[17] Our main results are based on a Poisson specification of the conditional mean:

$$\mathbb{E}(Y_{it}|W_{it}) = \exp\left(\beta \cdot X_{it} + \gamma \cdot AGE_{it} + \sum_{t=1998}^{2015} \mu_t \cdot \mathbb{1}_t \times AGE_{it} + \sum_g \omega_g \cdot \mathbb{1}_g \times AGE_{it}\right), \quad (3)$$

where *exp* denotes the exponential function, and $W_{it} = (X_{it}, Age_{it})$. This count specification has several advantages over the standard linear model, see e.g. Wooldridge (2002) or Cameron

[16]Before that date, these events were reported, but only a careful analysis of the full reports given by EDF to the ASN would allow to retrieve these events.

[17]Many of these variables are time constant.

and Trivedi (2013). Most importantly, the linear model might produce negative predictions for feasible values of the observed covariates, something we would like to preclude. Furthermore, the conditional quasi-MLE estimator of the Poisson model is fully robust to functional form misspecification, and if the model is specified correctly it is efficient. The interpretation of the coefficients is discussed in the results section.

Because the age of a reactor is necessarily correlated with the year of observation, the coefficients of AGE_{it} could be biased by the existence of some time-trend related to the efforts performed by the plant manager to curb the number of occurrences of safety events. To account for this relationship, we include in our specifications age-year interaction variables, defined as the product of the AGE variable and of a year dummy $\mathbb{1}_{YEAR=t}$. These variables will be noted $\mathbb{1}_t \times AGE$. Including these variables allows us to disentangle the effect of age, which is measured within each calendar year, from the effect of calendar time, which can be reconstructed as the variation over time of the coefficients associated with the age-year interaction variables. In order to measure the differences in declarations frequencies across the various technologies of reactors, we include group dummies in all our specifications ($\mathbb{1}_g$ in equation 3). Group dummies are interacted with the AGE variable in order to control for variations of the impact of age across capacity and conception groups.

We estimate the model described in equation 3 by running both Poisson and Negative Binomial models. One main advantage of the Negative Binomial regression is that it better fits over-dispersed data, which in our case is consistent with the descriptive evidence presented in figure 4 on page 12. A second advantage of the Negative Binomial estimator is that it is more efficient than the QMLE Poisson estimator in some cases, see Cameron and Trivedi (2013) for a discussion. In addition, our regressions allow for several specifications of the standard-errors, and our negative binomial regressions allow for linear (NB1) and quadratic (NB2) specifications of over-dispersion. The Akaike and Bayesian information criterion, and the Pearson statistics are used to select the regression model that best fits our data.

For each of these regression models, we estimate three baseline specifications (Specification 1, 2 and 3) which allow us to answer the questions which motivate this paper, e.g. whether the age influences nuclear safety, whether there are other significant drivers of nuclear safety, and whether the transparency of nuclear plant managers affect the declaration process. In specification 1, the dependant variable Y_{it} is defined as the yearly count of declaration of automatic shut-downs (ASD). The age of reactors is defined by the time elapsed since the beginning of the nuclear activity of their core.[18] and the group dummies are defined according

[18] Age could alternatively be measured with respect to other reference dates, such as the beginning of con-

to the three capacity groups presented in table 1.

In our second specification, we check whether the trends observed on automatic shut-downs can also be observed on another category of events characterized by perfect detection and declaration. This verification is important as automatic shut-downs have economic consequences for plant managers as they interrupt the production of electricity of a reactor. Hence, they provide incentives for plant managers to exert particular efforts to reduce their occurrences. In specification 2, we change the definition of the dependant variable, in order to consider the second categories of perfectly detected and declared events. Y_{it} is then defined as the number of safeguard events declared in reactor i during year t. All explanatory variables from specification 1 are left unchanged.

Our third specification aims to measure the importance of transparency in the declaration process. For this purpose, we change the definition of the dependant variable in order to relax our restriction to the set of perfectly detected and declared events. Y is then defined as the yearly count of safety significant events, ALL. We argue that transparency can be neglected if the results obtained under specification 3 match those obtained under the first two specifications. This claim will be discussed further in the results section.

Then, as several definitions can be given to some of the variables from equation 3, and that these definitions may influence the inferences drawn from our analysis, we propose two robustness checks on the definitions of our explanatory variables (Specification 4 and 5). Specification 4 differs from specification 1 in the definition of the AGE variable. With this second definition, we aim to capture the effects of the decade inspections undergone by each reactor, during which the main components of the reactors can be maintained or replaced. In specification 4, the age of the reactor is thus defined as the current decade of operation of the reactor, rounded to the upper decade. For instance, an 8 year-old reactor will be characterized by $AGE = 1$, while a 32 year-old reactor will be characterized by $AGE = 4$.

Finally, in specification 5, we change the definition of the group dummies in order to match the six conception groups introduced in table 1. This fifth specification allows us to check for heterogeneity among the different conceptions of reactors within each capacity group. These specifications are summarized in table 3.

As a remark, all our discussed specifications do not include individual fixed effects. The reason is that many of our explanatory variables are time constant and a within estimation would cancel them out. This is particularly problematic for the technology variable, which

struction, the first connection to the grid, or the beginning of commercial operation. The date of the first divergence is chosen as it best captures the amount of radiations received by the reactor's different systems.

Table 3: Model specifications

Nb.	Dep. Var.	Age definition	Group definition
1	ASD	AGE	Capacity
2	SFG	AGE	Capacity
3	ALL	AGE	Capacity
4	ASD	DECADE	Nominal Power
5	ASD	AGE	Capacity

is of main importance in the analysis. We have nevertheless run fixed effects regressions to obtain further evidence on the sign and significance of the AGE variable. The results are in line with the results obtained with our main specifications and we do not include them here. They are available upon demand.

4 Results and interpretation

4.1 Safety, age and technology

Calculations of the Pearson statistics for the Poisson regressions (using specification 1 and 2) allow to strongly reject the Poisson distribution for both automatic shut-downs and safeguard events. Moreover, the standard Akaike and Bayesian information criteria (AIC and BIC), computed using specification 1 under linear and quadratic over-dispersion, and different levels of clustering for the standard errors, provide strong support for the use of the negative binomial models with site-clustered standard errors and quadratic over-dispersion specification. These statistics are gathered in table 4. Results of the estimation of the negative binomial regression with quadratic over-dispersion and Site-clustered standard-errors for specifications 1 to 5 are presented in table 5. The results from the other regression models (Poisson and Negative Binomial with linear over-dispersion) are not reported. The following results are nonetheless robust to these alternative specifications.

Table 4: Test-statistics associated with the Poisson and NB regression models

Regression model	Over-dispersion	S-E clusters	Ln L	Pearson	AIC	BIC
Poisson	-	Reactor	-1455.5	1237.4	2965	3099
Poisson	-	Site	-1455.5	1237.4	2947	3036
Neg. Bin.	NB1	Reactor	-1447.4	-	2951	3089
Neg. Bin.	NB1	Site	-1447.4	-	2931	3020
Neg. Bin.	NB2	Reactor	-1446.5	-	2949	3088
Neg. Bin.	NB2	Site	-1446.5	-	2929	3018

Test-statistics based on model-1.

For a simple interpretation of the coefficients reported in table 5, the results of an OLS regression using specification 1 is provided (Specification 0). For the other specifications, the coefficients reported are obtained through a (non-linear) negative-binomial regression model, and can be interpreted using incidence rate ratios: given any explanatory variable X and its coefficient β_X, e^{β_X} represents the ratio of the expected counts of events obtained after and before a unit increase of X.[19] For instance, in equation 3, e^{γ} represents the (multiplicative) average effect of a unit increase in age on the expected number of occurrences of events in reactors belonging to the 900 MW tier, during year 1997. More generally, $e^{\gamma+\omega_g+\delta_t}$ represents the average effect of age on the expected number of occurrences of events within reactors of the group g during year t.

Consider first the results from specification (1). The main effect observed stems from the capacity group variables. The coefficient associated with the 1450 MW group is positive, significant, and shows that the reactors from the most recent capacity group declare on average a larger number of occurrences of automatic shut-downs. The coefficient associated with the age variable is also significant and positive, meaning that older reactors declare on average a larger mean number of automatic shut-downs. The significance and signs of the age-group interaction variables show that the effect of age is heterogeneous across capacity groups. More precisely, age seems to have a significantly larger effect on the reactors of the 900 MW capacity group.

The coefficients associated with the age-year interaction variables are significant, negative and globally decreasing. Thus, age affects nuclear power stations in a way that increases the probability of occurrences of automatic shut-downs, but the magnitude of this effect has been steadily decreasing over the last two decades. This improvement could be explained by a learning-by-doing effect at the fleet level, or by increased levels of safety care. Quantitatively, one additional year of age led to a 15% increase in the expected number of automatic shut-downs in the 900 MW group in 1997. This number falls down to 6% in 2014. For the reactors of the 1450 MW group, the trend is inverted: in 1997, the expected effect on automatic shut-downs of an additional year of age was almost 0, whereas in 2014, an additional year of age led to a reduction of this number by approximately 8%.

The coefficients associated with our control variables $SIZE$, $FOAS$ and $FOAK$ are insignificant, suggesting that the number of reactors within a plant has no significant effect on the occurrences of significant safety events. Likewise, neither being the first reactor within a plant ($FOAS = 1$) nor being the first reactor within a kind ($FOAK = 1$) seem to affect the

[19] When β_i is small for all i, β represents the vector of semi-elasticities of the dependant variable Y in the explanatory variables X. In addition, when coefficients are small and explanatory variables are included in logarithmic form, then β can be interpreted as a regular elasticity.

Table 5: Estimation results

	Baseline specifications				Robustness checks	
	(0)	(1)	(2)	(3)	(4)	(5)
VARIABLES	OLS	ASD	SFG	ALL	ASD	ASD
AGE	0.18***	0.14***	0.17***	0.016	0.52***	0.11***
1300 MW	0.95*	0.86**	1.35**	0.088	0.013	
1450 MW	2.95***	2.44***	2.64***	0.42	1.21***	
CP0						0.57
CP1						0.43
P4						1.41***
P'4						0.70*
N4						2.16***
1300×AGE	-0.033**	-0.030**	-0.017	0.0039	-0.060	
1450×AGE	-0.17***	-0.15***	-0.11**	0.0025	-0.90***	
CP0×AGE						-0.019
CP1×AGE						-0.023
P4×AGE						-0.056***
P'4×AGE						-0.037**
N4×AGE						-0.15***
$\mathbb{1}_{1998}$×AGE	-0.034	-0.025*	-0.025	-0.015**	-0.19*	-0.022
$\mathbb{1}_{1999}$×AGE	-0.049***	-0.036***	-0.035	0.011*	-0.25**	-0.031**
$\mathbb{1}_{2000}$×AGE	-0.055**	-0.041***	-0.048**	0.0066	-0.23*	-0.033*
$\mathbb{1}_{2001}$×AGE	-0.040*	-0.031**	-0.050**	0.00022	-0.14	-0.021
$\mathbb{1}_{2002}$×AGE	-0.052**	-0.039***	-0.044**	0.015**	-0.18	-0.028
$\mathbb{1}_{2003}$×AGE	-0.049**	-0.037***	-0.075***	0.020**	-0.14	-0.024
$\mathbb{1}_{2004}$×AGE	-0.075***	-0.056***	-0.10***	-0.0046	-0.29***	-0.041**
$\mathbb{1}_{2005}$×AGE	-0.081***	-0.061***	-0.073**	0.0041	-0.31***	-0.045**
$\mathbb{1}_{2006}$×AGE	-0.078***	-0.060***	-0.13***	0.0065	-0.28**	-0.042**
$\mathbb{1}_{2007}$×AGE	-0.086***	-0.066***	-0.11***	0.0057	-0.30***	-0.047**
$\mathbb{1}_{2008}$×AGE	-0.11***	-0.096***	-0.16***	0.0035	-0.56***	-0.077***
$\mathbb{1}_{2009}$×AGE	-0.096***	-0.073***	-0.098***	0.0081	-0.34**	-0.053**
$\mathbb{1}_{2010}$×AGE	-0.11***	-0.084***	-0.12***	0.00053	-0.42***	-0.062***
$\mathbb{1}_{2011}$×AGE	-0.11***	-0.084***	-0.11***	0.0057	-0.42***	-0.062***
$\mathbb{1}_{2012}$×AGE	-0.11***	-0.091***	-0.11***	0.0070	-0.46***	-0.068***
$\mathbb{1}_{2013}$×AGE	-0.11***	-0.085***	-0.13***	0.0058	-0.39***	-0.061**
$\mathbb{1}_{2014}$×AGE	-0.11***	-0.086***	-0.11***	-0.0020	-0.40***	-0.061**
$\mathbb{1}_{2015}$×AGE	-0.11***	-0.090***	-0.14***	-0.00071	-0.42***	-0.065***
SIZE	-0.061	-0.041	-0.21***	0.00097	-0.060	-0.026
FOAS	-0.16	-0.071	-0.32	0.016	0.043	0.012
FOAK	-0.10	-0.095	-0.041	-0.12	-0.25	-0.19
FOAS×AGE	-0.0037	-0.0055	0.00073	-0.0012	-0.086	-0.0066
FOAK×AGE	0.018	0.014	0.018	0.0092*	0.23**	0.016
Constant	-0.70	-1.50**	-2.46***	1.97***	-0.076	-1.22**
$ln(\alpha)$		-1.77***	-1.35**	-2.86***	-1.58***	-1.84***

1,100 Observations - Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

occurrences of significant safety events. This results suggests that the order of construction of similar reactors is not a significant driver of safety.

The signs and significance of the coefficients obtained under specification (4) are similar to those obtained under specification (1), suggesting that our results are robust to the definition of the AGE variable. The magnitude of the coefficient of the AGE variable is much larger than under specification (1), which is due to its new definition. Here, the coefficients captures the average effect of an additional ten years of age on the declarations of automatic shut-downs. Under specification (5), the results remain similar and significant, and show no sign of significant heterogeneity within capacity groups. Among the 900 MW reactors, results show that declarations in the CP0, CP1 and CP2 reactors are not significantly different. Likewise, even if the expected number of declarations from the reactors of conception P4 and P'4 are both significantly lower than in CP2 reactors, their two coefficients are not significantly different from one another.[20]

Results obtained under specification (2) are in line with those obtained under the first specification, confirming that similar trends can be observed when restricting the dataset to either automatic shut-downs or safeguard events. The magnitude of all effects are similar, supporting the hypothesis that the results obtained in specification (1) are not driven by specific efforts dedicated by plant managers toward automatic shut-downs. These two specifications thus support the hypothesis that age can be slightly detrimental to nuclear safety, that significant progresses have been achieved during the last twenty years, and that reactor technology is the most significant driver of nuclear safety observed in this study.

Finally, the signs and significance levels of our results are consistent across model robustness checks. Changing the negative-binomial specification from quadratic to linear over-dispersion, or replacing it by a Poisson model does not change our conclusion, nor does the change in the clustering level of the standard-errors.

4.2 Transparency

We now turn to the interpretation of specification (3), in which the dependent count variable is defined as the total number of events declared by nuclear reactors per year, without any restriction to specific types of events. Results of this regression are reported in table 5 on page 17. In this regression, the estimates of age and technology are not significant, although their signs are in the same direction. Similarly, the significance disappears for all other coefficients that were estimated in specifications (1) and (2). We now propose possible interpretations of

[20] An unreported t-test fails to reject the significance of the difference of these two coefficients.

this result.

Let the random variable T denote the type of an event, with $T = PDD$ being one of the possible types. Assume first that the effect of age and technology on the probability of event is the same for the full sample and the restricted set Θ_{PDD}:

$$\frac{\partial \mathbb{E}\{Y|W,C\}}{\partial W_j} = \frac{\partial \mathbb{E}\{Y|W,C,T=PDD\}}{\partial W_j}. \tag{A3}$$

Assumption (A3) is an external validity assumption. Then, if transparency plays no role in the declaration process, we should observe similar results for specifications (1), (2) and (3). We however fail to find significance in specification (3). Thus, either transparency does indeed bias the results, or assumption (A3) fails to hold.

However, assumption (A3) can be defended. First, comparison of regression (1) and (2) revealed no particular differences on the subsets of automatic shut-downs and safeguard events. Second, it can be argued that numerous investments in safety will have positive spillovers for all types of safety events. For instance, hiring skilled employees, investing in the training of safety engineers, or enhancing organizational practices are safety investments which will decrease the probabilities of occurrence of safety events, regardless of their types.

Provided assumption (A3) holds, our finding fails to reject the claim that transparency does not bias the analysis of the reports of nuclear safety significant events. Thus, neglecting the unobserved changes in declaration criteria, detection abilities, and rate of non-compliance with declaration criteria leads to significant under-estimations of the impact of age and technology on nuclear safety. This result contrasts the results of Rose (1990) who discarded the importance of pilots' subjectivity in airline incidents reports.

5 Conclusion

The takeaways from this paper are manifold. First, we describe a novel dataset obtained from the French Nuclear Safety Authority, encompassing over 19.000 significant safety events declared in the French fleet between 1973 and 2015. This dataset contains more events than previous datasets used for the assessments of the evolutions of nuclear safety. It also contains information of a better quality, as the declarations it gathers are verified by both nuclear plant managers and experts from the French safety regulator, whereas previous studies used press articles and academic publications to build their datasets.

Second, restricting this dataset to a subset of events characterized by perfect detection and declaration, we disentangle the impact of safety and that of transparency on the occurrences of safety significant incidents. Our results are consistent over both types of perfectly detected and declared events considered. This finding supports the hypothesis that our results are not driven by specific efforts exerted by plant managers and dedicated to one particular type of events. On the other hand, it appears that our results are not consistent when considering the unrestricted counts of safety events declared in each nuclear power reactor, which suggests that the level of transparency of plant managers during the declaration process introduces some bias which precludes the observation of safety variations when studying the complete set of safety events. This final result is in contrast with that of Rose (1990), who found that the subjectivity of airline pilots introduced no bias in the analysis of safety based on incident data.

Third, from a quantitative perspective, we observe that a significantly higher number of events are declared in older reactors. Yet, the magnitude of this effect is small when compared to the differences in the declarations that can be observed across reactor capacity groups. Moreover, the magnitude of this effect is decreasing over time. We first conclude from this analysis that the main observable driver of nuclear safety is technology, rather than age. In addition, even though age appears to be slightly detrimental to safety, it also appears that significant improvements in the management of ageing have been achieved over the period studied. Finally, the impact of age appears to be heterogeneous across technologies of reactors: the effect of age on safety has remained negative over the whole period observed for reactors of the 900 MW type, whereas it became positive for reactors of the 1450 MW type, suggesting the existence of safety improvements over the life of these reactors, a result in line with the findings of Rangel and Lévêque (2014) and Wheatley et al. (2016a,b).

Finally, we can derive some implications of these results for future nuclear policy. First, even though our study stands on French data, our results could have implications beyond the French fleet, as the pressurized water technology is the most widely used nuclear reactor technology, with 277 reactors operated on the planet according to the World-Nuclear Association. Second, our results suggest that focusing on the age of nuclear reactors when studying their safety, for instance by having large debates regarding the maximum lifespan of nuclear reactors, may be ill-advised. Instead, our paper suggests to focus on the particular weaknesses that characterize each reactor technology, as their impact on nuclear safety appears to be significant. In other words, we believe that setting a maximum lifespan irrespective of technological idiosyncrasies can be an inefficient policy, as it may lead to premature shut-downs of safe reactors, or prolonged operation of unsafe ones. A more appropriate policy suggested by our results would be to let competent nuclear safety authorities and nuclear operators decide when the investments required to maintain a reactor at an appropriate safety level become too high to justify its operation. Such a policy seems more appropriate, as it would lead to the identification and early closure of the less safe reactors.

Appendix 1: Proof of lemma 3.1

Proof of lemma 3.1. Note that

$$\mathbb{E}\{Y|W=w\} = \mathbb{E}\{\mathbb{E}\{Y|W=w,D,O\}\} = \mathbb{E}\{f(D,O)\}, \tag{4}$$

with $f(D,O) = \mathbb{E}\{Y|W=w,D,O\}$. In the following, we suppress the dependency on W for the sake of representational simplicity. It holds

$$\mathbb{E}\{f(D,O)\} = \sum_{d,o} f(d,o)\mathbb{P}\{D=d,O=o\}. \tag{5}$$

We observe that due to assumption A1, $\mathbb{E}\{Y|O=0\} = \mathbb{P}\{Y=1|O=0\} = 0$, and thus, under A1 and A2, equation (5) gives

$$\mathbb{E}\{Y|W=w\} = f(1,1)\mathbb{P}\{D=1,O=1\} = \mathbb{E}\{Y|W=w,D=1,O=1\}. \tag{6}$$

□

References

Cameron, A. C. and Trivedi, P. K. (2013). *Regression analysis of count data*, volume 53. Cambridge university press.

Faure, M. G. and Skogh, G. (1992). Compensation for damages caused by nuclear accidents: A convention as insurance. *The Geneva Papers on Risk and Insurance*, 17:499–513.

Hansson, I. and Skogh, G. (1987). Moral hazard and safety regulation. *The Geneva Papers on Risk and Insurance*, 12:132–144.

Hausman, C. (2014). Corporate incentives and nuclear safety. *American Economic Journal: Economic Policy*, 6(3):178–206.

Hirschberg, S., Burgherr, P., Spiekerman, G., and Dones, R. (2004). Severe accidents in the energy sector: comparative perspective. *Journal of Hazardous Materials*, 111(1):57–65.

Hofert, M. and Wüthrich, M. V. (2011). Statistical review of nuclear power accidents. *Asia-Pacific Journal of Risk and Insurance*, 7:1–13.

Laffont, J.-J. (1995). Regulation, moral hazard and insurance of environmental risks. *Journal of Public Economics*, pages 319–336.

Rangel, L. E. and Lévêque, F. (2014). How Fukushima Dai-ichi core meltdown changed the probability of nuclear accidents ? *Safety Science*, 64:90–98.

Rose, N. L. (1990). Profitability and product quality: Economic determinants of airline safety performance. *Journal of Political Economy*, 98(5):944–964.

Shavell, S. (1984). A model of the optimal use of liability and safety regulation. *Rand Journal of Economics*, 15:271–280.

Sovacool, B. K. (2008). The costs of failure: A preliminary assessment of major energy accidents,1907-2007. *Energy Policy*, 36(5):1802–1820.

Wheatley, S., Sovacool, B. K., and Sornette, D. (2016a). Of disasters and dragon kings: a statistical analysis of nuclear power incidents and accidents. *Risk analysis.*

Wheatley, S., Sovacool, B. K., and Sornette, D. (2016b). Reassessing the safety of nuclear power. *Energy Research & Social Science*, 15:96–100.

Wooldridge, J. (2002). Econometric analysis of cross section and panel data. *Cambridge, MA: MIT Press.*