The Trajectory of Psycho-Social Depression in Ukraine following the Chornobyl Nuclear Accident

Robert Alan Yaffee[1] , Thomas B. Borak[2], RoseMarie Perez Foster[3], Remi Frazier[2], Mariya Burdina[4], Victor Chtenguelov[6], and Gleb Prib[5]

[1] Silver School of Social Work, New York University, New York, N.Y.
[2] Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, Colorado
[3] Natural Hazards Center, University of Colorado, Boulder, Colorado
[4] Department of Economics, University of Colorado, Boulder, Colorado
[5] Institute for professional training of Public service of employment of Ukraine, Kiev, Ukraine
[6] Dept. of Social Work & Applied Psychology, Academy of Labor & Social Relations, Kiev, Ukraine

March 31, 2014

Keywords and phrases: nuclear incident, nuclear accident, psycho-social impact of Chornobyl, Chornobyl accident impact, nuclear incident impact, nuclear emergency planning, scenario forecasting, Chornobyl public health impact, confounding variable circumvention, Chornobyl impact on public health, Chernobyl psycho-social impact

Abstract

Our objectives were to examine predictive parameters of psychological impacts, resulting from the Chornobyl accident, on residents living in the oblasts of Kiev and Zhitomyr. We tested drivers for psycho-social depression based on estimates radiological dose received from radioactivity release during the accident and the perception of increased health effects associated with this radiation. To obtain a representative sample of individuals, we attached computer generated random numbers to area codes provided by the telephone company. In January 2009, Russia created an intervening crisis by interrupting supplies of natural gas to the Ukraine. We employed modified scenario forecasting to circumvent crisis effects that could otherwise undermine the internal validity of our study. State space methods were used to model and graph trajectories of psycho-social depression reported by male and female respondents. Results of the dose reconstruction process revealed that the dose received by this population was too low to identify pathological disease or injury. From our empirical analysis, we found that the psychological impacts of the nuclear incident stemmed from perceived risks, rather than actual exposure to radiation directly associated with the Chornobyl nuclear accident. Work funded by NSF HSD Grant 082 6983.

Contents

1 Acknowledgment

We would like to thank the National Science Foundation for supporting this research with HSD grant 082 6983 for modeling nuclear disaster risk and he Ministry of Health of Ukraine for cooperation in our study. We would also like to thank Sir David F. Hendry, Jurgen A. Doornik, Siem Jan Koopman, Andrew C. Harvey, Neil Ericsson, as well as Teresa Timberlake, David Corbett, and Noelia Januario of Timberlake Consultants, Ltd. for their support. Dr. Kathleen Tierney of the University of Colorado, Boulder, for her gracious support as well. Finally, we want to thank Jonathan D. Yaffee for his java program to locate all project data files for removal the personal identifying variables to assure anonymity of our respondents in compliance with Health Information Privacy Act requirements.

2 Introduction

Catastrophic incidents resulting in the release of radioactive materials receive unprecedented scrutiny by governments as well as individuals. These incidents might be accidental, (e.g., nuclear reactor malfunctions, improper disposal of medical instruments) or intentional, (e.g., nuclear weapons, radiation dispersal devices, improvised nuclear devices). The impact of these events can be severe, serious or incidental depending on the proximity to the event. Severe consequences are usually confined to the region from the epicenter extending out several kilometers. This region can be impacted by extreme heat, damaging blast, and exposure to very high levels of radiation. Serious conditions might extend as far as 30 to 50 kilometers and this geographical region is designated as an exclusion zone. Here, residual radiation originates from high concentrations of radioactivity deposited in the environment. This region often requires evacuation of all persons. Incidental regions can be global in extent; they occur where low levels of radioactivity can be identified above naturally occurring sources. In such regions, direct health effects are difficult to identify, but the concern for later effects persist for extended periods of time.

The aims of this work are focused on this extended region where there may be no indication of immediate health consequences, but psychological sequelae and perception of risk related to a nuclear incident can last for decades. As a basis for our model we use the nuclear accident at the nuclear power generating station near Chernobyl on April 26, 1986.

The accident at the Chernobyl nuclear power plant in 1986 was the most severe in the history of the nuclear power industry. Radioactive debris was expelled at the time of the initial thermal explosions and for the following 10 days during the ensuing graphite fires. It has been estimated that approximately 1017 Bq of ^{137}Caesium (^{137}Cs) was released. For comparison, this fallout is 10% of that released from all atmospheric nuclear weapons tests and about 10 times that of Fukushima [12].

The meteorological situation during the events was complex. Due to the long duration of the release and changing meteorological conditions, the radioactivity was dispersed in many different directions. The largest concentrations of radioactivity were deposited in Belarus, Ukraine and the Russian Federation. However, some radioactivity was effectively dispersed over much of Europe. The patterns of deposition were irregular with Hot Spots adjacent to Cold Spots. Data from extensive measurements were compiled into maps in the Atlas of Cesium Deposition on Europe following the Chernobyl Accident published by the European Commission [12].

3 Research objectives

1. To empirically test whether the dose to individuals from external sources of radiation released during the accident was a significant explanatory variable in describing the temporal patterns of depression in the population of residing in the oblasts of Kiev and Zhitomyr in Ukraine.

2. To empirically test whether the perceived health risk associated with radiation from the Chernobyl accident was a significant explanatory variable in describing the temporal patterns of depression in the population of residing in the oblasts of Kiev and Zhitomyr in Ukraine.

3. To develop, formulate, predictively validate, and assess a forecasting model to project the level of depression to be expected after a nuclear incident. To devise a forecasting protocol that can circumvent the impact of major non-Chornobyl-related intervening variables that could confound our analysis of the trajectory of self-reported depression following the accident at Chornobyl

4 Literature review

4.1 Issues with epidemiological case-control studies

Epidemiological case-control studies purport to compare Chornobyl radiation-exposed segments of the population to presumably unexposed reference groups. The exposed groups included cleanup-workers, some of who had leukemia and solid cancers.. They also included studies of young children whose thyroid cancers were associated with ^{131}Iodine intake. Other studies focused on portions of the 116,000 evacuees from the exclusionary zone of 30 km around the power plant as well as immigrants seeking refuge from adverse environmental effects. Observational studies without proper empirical calibration have been found to fail to correct for systematic error induced by selection bias, misclassification, and residual confounding, resulting in spurious significance assessments. Without specific correction for such systematic error, these biases are believed to be "intrinsic to observations studies in general [46, 209,210,216]." In other words, without specification of a random sampling protocol, and randomized neutralization of such biases, the external validity of these studies may remain in question [24].

4.2 Primary psycho-social symptomatology

The Chernobyl Forum Report of the 20th anniversary of the Chernobyl accident concluded the most significant public health consequences were social-psychological [7]. It is not surprising that the first responders and cleanup workers who had the greatest exposure to radiation would exhibit elevated rates of depression, anxiety, and PTSD for years thereafter. In a review of the psychological consequences of the Chornobyl accident after 25 years, Bromet, Havenaar, and Guey (2011) not only affirm that conclusion, they suggest that the lives of evacuees were disrupted by social uprooting, relocation, social discrimination, and stigmatization [8, 297-298]. Moreover, they maintain that "the Chornobyl disaster encompassed a vast array of physical and psychosocial exposures that are all but impossible to disentangle from the general turmoil that followed the collapse of the Soviet Union in 1991 [8, 298]." However, in their review of studies conducted, the authors point to only one all study that included a representative sample as a reference group [8, 301-303]. In this paper, we focus on the depression symptoms reported by respondents, all of whom were randomly selected.

4.3 Principal confounding variables

Bromet (2012) noted that after the collapse of the Soviet Union in 1991, there was a vast array of events that are almost impossible to disentangle from those following Chornobyl [6]. Some of the most salient events appear to have followed the November 2004 Orange Revolution in Ukraine, [9, 7]. The Russian natural gas trading dispute in 2005 resulted in the brief natural gas shut-off in winter (January) of 2006, the 2007 Ukrainian political crisis involving a power struggle between the President and Parliament. The Great Recession emerged in September 2008 and in winter, 2009, the Russians shut-off all natural gas flowing to the Ukraine over a dispute about prices and unpaid bills [4]. Most of the factories in Ukraine were forced to cease production, which imposed much hardship on the Ukrainian economy. The effects these Russian actions had on Ukrainian anxiety and depression were very likely to have been substantial. Unless these intervening and confounding effects can be distinguished, attempts to relate psychological sequela to the Chornobyl accident may be invalidated. Most earlier studies do not show how these effects were disentangled. We take explicit precautions to avoid these confounding impacts on our measure of reported depression following Chornobyl.

5 Research design

In this paper, we focus on the particular psychological symptom of depression, and we will address other symptoms in future papers. We test whether external dose of radiation was a significant explanatory variable in the explanation of the psychological symptom under consideration on the part of the Ukrainian public in Kiev and Zhitomyr oblasts. We test whether perceived Chornobyl–related health risk was a significant variable in the explanation of psychosocial depression after the Chornobyl nuclear accident. We find a way of circumventing major confounding events near the end of our data.

5.1 Representative sampling

Our study differs from those cited above by using only random selection to obtain a representative sample. Phone numbers were randomly generated, attached to area codes supplied by the Ukrainian telephone company, and called to make an appointment for an interview. Up to four calls were placed to selected numbers in the two oblasts. In total, 702 respondents, consisting of 339 males and 363 females, to comprise the representative sample of the residents in Kiev and Zhitomyr oblasts. Kiev oblast is where the Chornobyl nuclear power generating plant was located, and Zhitomyr oblast is just west of and adjacent to Kiev oblast. Before the data for the interview was uploaded, a independent audit of the interview was conducted to confirm that all answers were given freely and voluntarily. Once confirmation was provided, the responses were uploaded.

5.2 Interview and survey format

The survey was administered in person by the interviewer at the home of the respondent. The format of the interview was that of a retrospective panel analysis. The survey was organized in this manner to facilitate recall and to reduce any bias. Although the survey was organized in this manner the initial data format was stored in a rectangular wide format. It was reorganized for panel data analysis.

Because no one could ethically plan such a disaster, all questions had to be posed retrospectively, lest the nature of the public response be lost . Questions were posed in the context of easy temporal markers to facilitate recall and to minimize recall bias. The time frame of recollection consisted of a prelude period before the Chornobyl accident, and three easily recalled periods (waves) of recollection. The prelude period extended from January 1, 1980 until April 26, 1986, the date of the Chornobyl accident. Wave one extended from the date of the disaster until the end of the same year. Wave two began January 1, 1987 and extended through 1996. Wave three began in January 1, 1987 and extended through the end of 2009 for dose reconstruction. The interviews for the general sample of 702 were conducted from 2009 through 2010, and a few in 2011. For the purpose of a panel data analysis, the dataset of more than 2500 items was reorganized into a panel dataset.

Some of the responses were measured in accordance with the change in level of a self-reported symptom. When these responses were collapsed by year, they provided us with a time series dataset which would facilitate our analysis of historical trajectories and trends in the psychological symptomatology. It is this dataset that is used for the analysis of the historical male mean and female mean response for self-reported depression level, which constitutes the subject of this paper. Items from these datasets were merged into one time series dataset for the purpose of analysis for this paper.

Because this analysis is a psycho- socio-medical analysis, we were interested in the responses of males and females. It is obvious that the responses will differ between them. Because we are interested in the nature of those differences, we perform the same but separate analysis for male and female respondents. After the data were cleaned and prior to statistical analysis, all indicators of specific personal identity were removed from the datasets to protect the confidentiality of the respondents.

5.3 Measures and indices

The survey instrument was a questionnaire that was translated into Russian and back-translated to assure congruency of meaning. This questionnaire was administered by personal interview that was pre-arranged by appointment with the respondent. After the interview a separate check was conducted to assure that all item answers were voluntary. When confirmation of this volition was obtained, the data were uploaded and input by the Vovici corporation into a file, which was converted to Stata for data management.

Our endogenous variable is constructed in a manner that it can be transformed into a time series. Had we used the conventional definition of depression from the Beck, Hamilton, BSI, or other standard composite battery, they would not have been readily amenable to transformation into a time series with enough repeated observations to have been useful. We construct our endogenous variables of psycho-social depression (variables `maledep` for males and for `femdep` for females) by annually computing the sample mean self-reported level of depression, measured in units of percent (between 0 and 100), after we drop the lower five percent to be sure this was not to be confused with a mood of sadness.

If we make the working assumption that the recollection of our respondents was sufficiently accurate, because we focused on significant changes over time of depression and because we are examining the effects of a major incident, for which there was almost no history to use as a basis for emergency planning, we use these annual averages of recalled depression as the basis of our analysis. Depression scores were computed from 1980 to the present. Moreover, we forecast from 2005 forward to show what the psychological effect of Chornobyl accident was, rather than to include events which would cause psychological depression by themselves and therefore confound the relatively isolated recollection of the effects of the accident at Chornobyl.

The average number of illnesses reported per wave–namely, `millw` for men and `fillw` for women, were also computed from 1980 to the time of the interview (Figure 2).

If we forecast from 2005 onward, we have an estimate of the Chornobyl accident impact. Any level of depression exceeding the upper predictive confidence limit of our forecast would be a level attributable to the other intervening events. The probability density fan around the forecast can be used to define the different levels of confidence. We use a one forecast standard error to define the probability density around the forecast. It highlights the difference between our forecast and the actual level of depression realized over the forecast horizon.

We can include a measure of the time-varying age of the population as it steadily increases each year, but we have to transform this variable by differencing the natural log of the to render it covariance stationary prior to using it in the model. The variables are called `dlmaleage` for males and `dlfemage` for females. If we do not use such a transformation, we unnecessarily risk spurious results, Granger and Newbold (1974) suggested [21]. However, as such, the age variable turned out to be non-statistically significant in the initial model for both males and females, for which reason it was dropped from the final model. Because these variables comprise diagonal lines prior to transformation and flat lines following transformation it is not surprising to find them unrelated to change in the endogenous variables, for which reason we do not include it in Figure 2.

We include the effects of dose from external penetrating gamma rays (variables `mavgcumdose`) for males and `favgcumdose` for females) on the psychological symptoms. We test the exposure to radiation from the radioactivity released during the Chornobyl accident influences the temporal patterns of self-reported depression. The unit of measurement of reconstructed external dose is the milliSievert (mSv). The reported results from dose reconstruction exclude contributions from natural background radioactivity or cosmic rays.

A process was developed to reconstruct the dose from penetrating gamma rays emitted by radioactivity deposited on the ground to each individual in the survey as a function of time. The radiation source term was obtained from the *Comprehensive Atlas of Caesium Deposition on Europe after the Chernobyl Accident (1998)* [12]. This document includes maps showing ^{137}Cs concentration across Europe, presented in equal-area Lambert oblique azimuthal projections.

The electronic version of this Atlas includes each map plate stored in a vector graphics format with multiple layers of information. One of these layers shows isolines representing intervals of equal ^{137}Cs deposition at the time of the accident; an overlaid layer provides a labeled grid corresponding to intersections of latitude and longitude (this is properly referred to as the conjugate graticule).

Software was developed to recover the contour color that specifies the ^{137}Cs concentration at a specified

Table 1: **Cumulative Dose (mSv)**

External Dose Summary Measure	**12/31/1986**	**12/31/1996**	**12/31/2009**
Lowest value of External Dose received by an individual	0.0074	0.036	0.047
Largest value of External Dose received by an individual	28.0	30.0	31.0
95th % Quantile of External Dose received by the sample	[0.037 – 1.4]	[0.14 – 3.4]	[0.19 – 4.4]
Average value of External Dose received by the sample	0.38	0.93	1.2
Standard Deviation of External Dose received by the sample	1.2	2.0	2.2
Median value of External Dose received by the sample	0.28	0.69	0.91
Estimated Average value of External Dose from Natural Background	0.33	5.3	12.0

latitude and longitude. This was accomplished by using the intersections of the conjugate graticule as a guide to define a transformation from the original Lambert projection into an equirectangular projection. This transformation was then applied to the map layer which showed ^{137}Cs concentration, which allowed the ^{137}Cs concentration maps to be loaded into a geostatistical database. Conversion tables between published isoline colors and indicated ^{137}Cs concentration were produced. Latitude and longitude coordinates could then be submitted to the geostatistical database in order to recover the ^{137}Cs concentration at an arbitrary location. Where map plates published in the Atlas overlapped, the ^{137}Cs concentration was taken from the map with the most spatial detail; if a location was submitted to the geostatistical database which had no corresponding map data, the closest available ^{137}Cs concentration was used.

A model was created to determine the dose rate at an arbitrary time t for any individual in the study. This model is based on the following sequence of factors: 1) ^{137}Cs concentration at a location (Lat. Long.) at the time of the accident,$C(t_0)$ [12], 2) ^{137}Cs concentration, at time, t, based on decay, soil mixing and weathering, C(t) [36], 3) Kerma rate to air, K(t), from penetrating gamma rays emitted by all radioisotopes, normalized to the Cesium concentration C(t) [35], 4) Conversion from kerma in air-to-dose in person, as a function of age, at time t [12], [36], [35], [45], [28], & [29]; 5) Modifying factors for time spent outdoors based on occupation and age. [35]; and 6) Shielding factors based on residency indoors and typical construction [35].

The data are integrated and presented as the annual dose rate received by each individual in units of mSv/year.

Figure 1 shows the results of the dose reconstruction for males and females in terms of annual dose rate, milliSieverts per year, (mSv/y) and time integrated cumulative dose (mSv), and Table 1 listed the descriptive statistics for the cumulative dose.

To assess perceived Chornobyl-related health risk, we constructed a perceived Chornobyl-related health risk index `crhrw#` consisting of an index of three items, measured separately in each wave of the study. From 1980 to the date of Chornobyl accident, we called the prelude period. Wave one covered the time from the date of the Chornobyl accident, April 26, 1986, through the end of that year, December 31, 1986. Wave two covered the period of January 1, 1987 through December 31, 1996. Wave three covered the time since

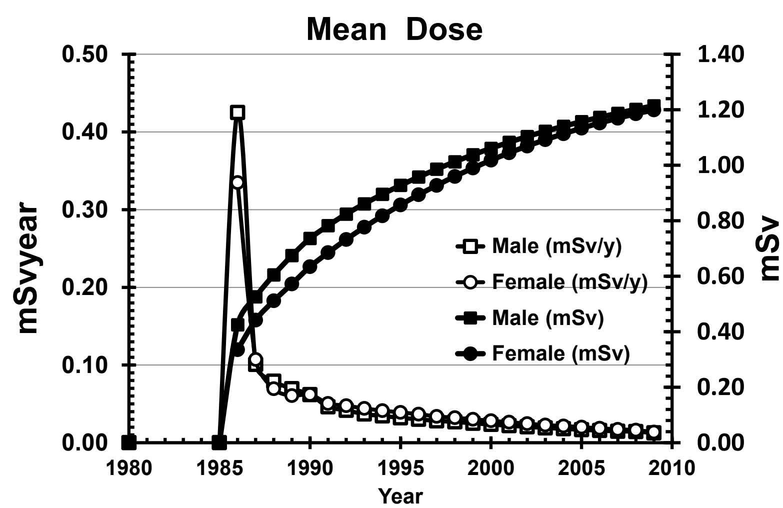

Figure 1: Cumulative external dose and differential external dose excluding natural background and voluntary medical radiation

Table 2: Perceived Chornobyl-related health risk `crhrw`(#) Cronbach (α) reliability

Chornobyl- related health risk index	**male α**	**female α**
Wave 1 : crhrw1	0.822	0.761
Wave 2 : crhrw2	0.835	0.796
Wave 3 : crhrw3	0.841	0.818

January 1, 1997 until the end of 2009. The interviews for the general sample began in 2009 and ended in 2011.

The variables averaged in the index were 1) How much in percent the respondent believed Chornobyl accident affected his or her health, 2) How much in percent respondent believed that Chornobyl accident affected his or her family's health, and 3) the percent of belief in the statement that in Kiev/Zhitomyr oblasts, most human cancer cases are known to be caused by radiation. We justified the combining these items by the scale α reliability exceeding the recommended Cronbach's alpha reliability of of 0.70 for scale inclusion, as indicated in Table 2 [42]. Thus, our averaged index of perceived Chornobyl accident health risk covers the self, the family, and the community in level and scope.

When the Chornobyl related health index was broken down by gender, we called the male variable, `mrpre2,` and the female variable, `frpre2`, displayed in the initial models of the following tables.

The third major candidate explanatory variable was the recollection of the respondent of the number of illnesses experienced during each wave. We tested all of these variables in the initial gender-specific model.

All procedures applied were approved by the Institutional Review Board at the University of Colorado, Boulder, before they were implemented.

5.4 Confounding variable circumvention

Ukrainian history is replete with events after 2005 that are likely to confound any psychological analysis of depression, anxiety, or even PTSD. They entail gas supply and price disputes followed by natural gas cut-offs in the winters of 2006 and 2009. They include political crises in 2007 over the dismissal of Parliament, scheduling and paying for new elections, and commotion over meddling with Constitutional Court membership [4]. Although we do not want to get into a detailed discussion of the relative impact of this new array of sources of discontent, we need a way to circumventing their impact on the remnants of self-reported depression from the impact Chornobyl on depression. We observe a surge in depression during this 2006 through 2010 period of difficulty that is larger than even the impact of Chornobyl. In 2006, there was arctic winter blast that killed 53 people and later that year there was an outbreak in the Crimean part of Ukraine that hospitalized approximately 3000 people. Without a means of avoiding the impact of these intervening events, it may become impossible to distinguish the impact of Chornobyl on the impact of these other events, thereby compromising the internal validity of any retrospective study of the psychological impact of those events. Although we recognize that this method may not be a panacea to this problem of disentangling other more minor sources of depression with that stemming from Chornobyl, it should suffice to historically distinguish the major challenges to the recollection of what came from Chornobyl. We attempt to circumvent the intrusion of these intervening impacts by reverting to an earlier point of forecast origin, and forecasting remaining depression over this time period with a type of conditional forecast of an alternative scenario.

6 Statistical methods

6.1 State space model selection

We needed a method that could develop an explanatory model that could be used with relatively small data sets. We needed a model which we could use to test key explanatory variables. We needed a model that could perform both out-of-estimation-sample *ex post* and beyond the end of the actual data *ex ante* forecasts. We also needed a method that could incorporate mixed frequency variables along with dummy impact indicators to identify and measure a latent state or structure. We needed a model that could update such a latent structure and smooth it for signal extraction. Because exponential smoothers has been shown to forecast well with such datasets and because state space models have been interpreted as an elaboration of exponential smoothing [25]., and because some of our data are sampled at different frequencies, we decided that this method, which had been applied to mixed frequency data to analyze a latent state vector, could be applied properly to our data. In sum, we needed a statistical model with which we could do hypothesis testing, model-building, and forecasting with mixed frequency data.

The statistical method of choice was state space (unobserved components) modeling, originally developed by Richard Kalman in 1960. This technique is used for tracking rocket trajectories, after Dr. Stanley F. Schmidt of the NASA-Ames Research Center modified the Kalman Filter so it could be used for navigation of Apollo 8 flights to the moon [27], [5, xiii, 13, 431][43, xv, 206, 213].

After summarizing this method, we will organize our statistical analysis as follows. In all cases we visually review the shape of the reported depression trajectory depicted in time series plots. We will present an initial model which will determine which unobserved components– for example, level and slope– and which explanatory variables significantly contribute to the explanatory model from which we can reliably forecast . In the initial model , we test explanatory variables in our research questions to empirically ascertain the answers those questions. Therefore, we test cumulative exposure to ^{137}Cs as an indicator of cumulative external dose, perceived Chornobyl-related health risk, along with other variables or event indicators that significantly improve the explanation of annual average of self-reported depression.

Second, we prune from this initial model all statistically nonsignificant variables, components, and event intervention dummy variables not needed for the trimmed model. We trim the model to obtain the most parsimonious, encompassing model that optimally explains and predicts the reported levels of post-accidental depression.

Third, we subject these components, variables, and event dummy variables to a battery of misspecification tests. We test the model for validation with the misspecification tests. Therefore we test the model residuals for independence, normality, homogeneity, and autocorrelation. We use standardized residuals to examine serial correlation or lack of independence. Independence is checked with the Box-Ljung Q statistic; residual normality is tested with the Bowman-Shenton test, homogeneity is checked with a Chow breakpoint and a CUSUM squared test, and residual autocorrelation is tested graphically with correlograms. We check the CUSUM test for parameter stability, along with index plots of the auxiliary standardized residuals to identify outliers or level shifts. We combine the tests for normality, skewness, and kurtosis to obtain a quasi-Jarque-Bera assessment of the behavior of the residuals. Although we do not display all of this output here, we indicate these results in a misspecification test summary table. If the variables remain statistically non-significant and pass the misspecification tests, they are retained in our final trimmed model.

Fourth, we test for predictive validation of the model. Merely because a model fulfills the assumptions for model validity does not mean that it has predictive validity. If the model is over-fit, over-parameterized, or over-specified, part of the parameterization would be modeling the noise as well as the signal. The result would be a high goodness-of-fit coupled with a poor forecast. In this section, we evaluate the model for over-fitting or over-parameterization. The process of predictive validation entails an *ex post* forecast evaluation, in which we compare the observed data in the validation segment with the forecast over that segment. If the results indicate no significant difference between the observed and predicted within this validation segment, the predictive validation will have been fulfilled. Passing this test provides a license for *ex ante* forecasting beyond the end of the data.

Fifth, we project employ iterated projections, sometimes called one-step ahead forecasts, of depression from 2005 to the end of our forecasting horizon in 2010. For this purpose, we employ Root Mean Square Forecast Error (RMSFE), Mean Square Forecast Error (MSFE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) to confirm the accuracy of our forecasts against the remainder of our data.

Sixth, we perform an evaluation of the that forecast. By using the hold-out or validation segment of the data for this, we have a reliable basis of comparison for our ex ante forecast evaluation. The criteria used for this evaluation are the error (actual - forecast), the root mean square error (the standard error of the forecast), the root mean square percentage error (which takes only positive values), and to minimize outlier distortion we also use the mean absolute error, and the mean absolute percentage error.

6.2 The date of forecast origin

The measures we take to overcome this challenge to the internal validity of the study may appear novel, but not radically new. We forecast from a point of origin at the beginning of the year of 2005 so that our forecasts, which project the impact of Chornobyl on depression, are not compromised by challenges of subsequent major confounding or intervening variables that occur after 2004. In this way, we displace large confounding or intervening impacts that could confound our analysis of psycho-social depression. Redefinitions of starting positions are commonplace in forecast evaluation.

Tashman (2000) has advocated flexibility in the position of the forecast origin for forecast evaluation to overcome corruption of evaluation by end effects. With rolling origin forecasts, the forecast accuracies are computed for various forecast horizons and then averaged to obtain an omnibus measure for the method [48, 439].

We combine an element of this approach with that of a scenario forecast, predicated upon a spectrum of variable conditions fundamental to the forecast outcome. In 1950s and 60s, Herman Kahn, as a Rand Corporation military strategist, conceived of the "ladder of military escalation," for defense planning [30] Michel Godet in 1982 suggested forecasting different futures by problem formulation, key variable searches, formulation of strategic stakes and objectives, and the formulation of scenarios as probabilistic road maps to alternative futures [20]. Sensitivity analysis entails a variety of what-if questions and answers to reveal fundamental variability of a model. Ed Leamer's extreme bounds analysis (1983) is an example of this kind of analysis [34]. Pierre Wack (1985) explained the Royal Dutch Shell application of multiple scenario analysis as a planning tool for the oil crisis of 1972-73 [50]. The alternative future that we consider is what would have occurred had the process not been impacted by confounding variables possibly behind the subsequent observed surge in the symptom under consideration. The statistical technique we employ is structural time series analysis owing to its applicability in longitudinal impact analysis.

We find that if we use 2005 as a point of forecast origin, our model can explain the trajectories of reported depression with predictive validity and it can avoid the political and economic turmoil of 2006 through 2009. At the beginning of both of these years, the flow of natural gas to the Ukraine was cut off, which caused considerable anxiety and depression on the part of those already suffering from it.

6.3 Structural time series analysis

Although we used Stata for general data management, we selected the OxMetrics Suite in general, and STAMP, in particular, by Siem Jan Koopman, Andrew C. Harvey, Jurgen Doornik,and Neil Shepherd, for our analysis. This package has a track record of superb performance and accuracy when applied to time series analysis [52] With state space models, we focus on latent components, explanatory variables, and event interventions of the data generating processes we are analyzing. As we identify the model elements, we check them for model fit, component integrity, historical accuracy, hypothesis testing, predictive validity, and forecast accuracy, during the model building process. The state space model consists of a system of equations including a transition and a measurement equation. The state space models are named according to the principal components or factors in the model. In this explanation, we follow the notation of Koopman, Shephard, and Doornik in SsfPack 3.0 [33].

6.4 Transition equation

The transition equation models unobserved time series components as they evolve over time. The transition process of moving a state (of latent structure, explanatory variables, and event intervention indicators) from one point in time to the next is that of a first-order autoregressive projection plus a regression on the shock or innovation. This transition, followed by an iteration of the measurement phase, an augmented factor analysis, is called Kalman filtering [1, 36]. It is a process of the state (latent structure) being iteratively projected and redefined from one time period to the next. It is a process of Bayesian sequential updating of the mean and variance by taking weighted averages of the current and past values to obtain expected values of those parameters.

$$\alpha_{t+1} = T\alpha_t + H_t\epsilon_t \tag{1}$$

where α_t = a m x 1 state vector, where m = the number of rows (components + variables + dummy indicators) in the state vector, T = an m x m matrix of transition coefficients, and H_t = an m x r transition error vector, and ϵ_t is an r x 1 error vector.

6.5 Measurement equation

The measurement equation defines a principal components or factor analysis to define the latent structure (of components or factors). At each period of time, a factor analysis uses the observed variables to rederive the latent structure. Thus the latent state (factor structure) is observed to be redefined.

The factor analysis is defined by the measurement Equation 28. The factors consist of a state vector, shown comprised of components that have a sufficiently high signal-to-noise (q) ratio, where the noise comprises the irregular component in the time series. The measurement equation can model a level, a slope, a seasonality, and a noise component to comprise- the parameters that are included in the state vector for updating. The measurement equation defines the relationship between the observed variable, y_t and the latent state vector, α_t, and the error vector G_t in an abbreviated form, such as

$$y_t = Z_t\alpha_t + G_t\epsilon_t \tag{2}$$

where our observed endogenous variable, y_t is an N x 1 vector (N=the number of variables) is a product of a N x m factor loading matrix, Z_t and the G_t, a N x r measurement error squared vector at that time.

6.6 Kalman filter

In this presentation of the Kalman filter, we follow the work of Koopman, Shephard, and Doornik (2008) and their notation [33, 7-49]. From these two equations, we are able to compute four system matrices: Φ_t, an (mxN) by m matrix , Ω_t, an$(m+N)$ by $(m+N)$ matrix, u_t, an $(m+N)$ by $(r+r)$ matrix, and an (m + 1) by m matrix of initial conditions, where N = the number of variables, n = the number of observations, m = the number of rows in the state vector, and r = the number of rows in the error vectors (G$_t$)$or(H_t)$.

$$\Phi_t = \begin{bmatrix} T_t \\ Z_t \end{bmatrix} \quad \Omega_t = \begin{bmatrix} H_t H_t' & H_t G_t' \\ G_t H_t' & G_t G_t' \end{bmatrix} \quad u_t = \begin{bmatrix} H_t \\ G_t \end{bmatrix} \epsilon_t \tag{3}$$

Therefore, we can stack the transition equation atop of the measurement equation and process the system as follows:

$$\begin{bmatrix} \alpha_{t+1} \\ y_t \end{bmatrix} = \Phi_t \alpha_t + u_t \epsilon_t \tag{4}$$

Because we assume mean centering,

$$\nu_t = y_t - Z_t \alpha_1 \tag{5}$$

We can compute the variance of ν_t to obtain

$$Var(\nu_t) = F_t = P_t + \sigma_G^2 \tag{6}$$

$$F_t = Z_t P_t Z_t' + G_t G_t' \tag{7}$$

where F_t = the total variance, Z_t = an N x m the factor loading matrix of observations by variables, P_t = the m x m covariance between the observable variable and the state vector (variance of the state vector), and G_t = an r x 1 measurement error vector, and GG' = the r x r specific error variance matrix.

First starting values are obtained for the state vector and its variance, and if the components of the state vector have been mean-centered, they can be set to zero. The values of the initial variance P_1 is set to a very large number.

$$\alpha_0 = (\alpha_1, P_1) \tag{8}$$

The innovation is computed.

$$\nu_t = y_t - Z_t \alpha_1 \tag{9}$$

From this error, we can compute its variance

$$F_t = Z_t P_t Z_t' + G_t G_t' \tag{10}$$

To obtain the coefficient of the Kalman gain, K_t, the following formula are used.

$$M_t = (T_t P_t Z_t' + H_t G_t') \tag{11}$$

$$K_t = M_t F^{-1} \tag{12}$$

from which we can project the state vector

$$\alpha_{t+1} = T_t \alpha_t + K_t \nu_t \tag{13}$$

and its variance

$$P_{t+1} = P_t(1 - K_t) + H_t H_t' \tag{14}$$

$$\begin{bmatrix} P_{t+1} & M_t \\ M_t' & F_t \end{bmatrix} = \Phi_t P_t \Phi_t' + \Omega_t \tag{15}$$

after which we update until we reach the end of the series, with

$$\nu_t = y_t - \hat{y}_t \tag{16}$$
$$\alpha_{t+1} = \bar{\alpha} + K_t \nu_t \tag{17}$$
$$P_{t+1} = \bar{P}_{t+1} - K_t M_t' \tag{18}$$

For a detailed explanation of the Kalman filter recursions, Harvey (1989, 104-113) and Durbin and Koopman (2000, 12) also are recommended.

7 Local level model

There are two state space models that we find most useful for estimating our models. The first of these version is the local level model plus regressors and event dummy variables. We will discuss the local level model plus noise first.

7.1 Measurement equation

In this model, the measurement equation defines a simple process consisting only of a local (which can change from one time period to the next) level, plus an error (irregular component) vector, G_t. The local level, μ_t, in time series is sometimes called a drift parameter. G_t is the irregular (error) component. We can formulate the measurement equation for the local level model as

$$y_t = \mu_t + G_t \quad G_t \sim NID(0, \sigma^2_{G_t}) \quad for\ t = 1, ..., T \tag{19}$$

where y_t = the observed variable, μ_t = local level component, and G_t = the measurement error.

7.2 Transition equation

The transition equation for the local level model could be expressed as

$$\mu_{t+1} = \mu_t + H_t \quad H_t \sim NID(0, \sigma^2_{H_t}) \quad for\ t = 1, ..., T \tag{20}$$

where μ_{t+1} = a projection of the local level component ahead one time period into the future, μ_t = a local level component at time = t, and H_t= the transition equation error, moving along time path t = 1, ..., T. The transition error is distributed normally and independently with mean = 0 and variance = $\sigma^2_{H_t}$.

7.3 Initialization, estimation, and smoothing

The Kalman filter is a Bayesian updating sequential process that requires starting values. The sequential updating is done with a weighted average, weighted according to the precision (inverse of the variance) of the estimates. With no prior information, a non-informative prior variance must be applied to the to construct the weights. A very large variance approximating is applied to construct a tiny weight for a prior distribution; hence the term a diffuse prior. The process is initialized with a by applying starting values for the Kalman filter. The initial conditions are specified in an (m + 1) by m matrix. When not otherwise specified, it is assumed that the initial conditions are fully diffuse, such that

$$\sum = \begin{bmatrix} P_1 \\ \alpha_1 \end{bmatrix} \tag{21}$$

where $\alpha_1 \sim N(0, \kappa I)$ and $P_1 = P_* + \kappa P_\infty$ with $P_* = 0\ and\ P_\infty = I,\ with \quad \kappa \to \infty$

During estimation with maximum likelihood, the filtering uses the sequential updating with a weighted average of the previous and current values to obtain the predicted values of the mean and variance of the parameter. When the parameter estimates converge to a steady state and the results are reported.

After the filtering has been completed and all of the data are collected, the smoothing of the signal can take place, using all of the data simultaneously, in the reverse direction. From the smoothing all the data, we extract signal from the noise. Because this algorithm is based on a random walk plus noise, it works well with nonstationary data, provided that an augmented Kalman filter, which we will explain shortly, is used. We will also discuss adding the regressors and intervention indicators after we discuss adding a local trend component in the next section.

8 Local linear trend model

A local linear trend model consists of a local level, μ_t, and a slope component, β_t, to an irregular (error) component, G_t. The local level and local slope parameter are together referred to as the trend, although the slope parameter alone is often called the trend term in a simple OLS regression.

8.1 Measurement equation

The measurement model of this local linear trend model can be expressed as

$$y_t = \mu_t + \beta_t + G_t \tag{22}$$

8.2 Transition equation

The local linear trend model has a transition model comprising updating equations for both the local level and the slope components, such that

$$\mu_{t+1} = \mu_t + \beta_t + H_t \quad H_t \sim NID(0, \sigma^2_{H_t}) \tag{23}$$

The slope parameter could be fixed (if the error term were set to equal zero) or it could be random (stochastic) if the slope error term were nonzero and significant. If the q of the slope parameter is not significant, the slope is generally automatically dropped from the model. If the slope does not contribute to the model we could dispense with the slope term, (β_t), altogether and rely on the local level model. However, if the slope signal to noise ratio, q_β, contributes significantly to the process, we would keep the β term in the measurement equation and include a slope updating or transition equation

$$\beta_{t+1} = \beta_t + H_t \quad \zeta_t \sim NID(0, \sigma^2_{H_t}) \tag{24}$$

8.3 The state vector

We can handle such multiple equations systems by vectorizing or stacking their endogenous components in a state vector. For example, a local linear trend model would have two parameters, the local level and the slope, which we could stack in a state vector, α_t as follows

$$\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} \tag{25}$$

We could then project our state vector along a trajectory in accordance with the state space recursions.

$$\alpha_{t+1} = T\alpha_t + G_t \tag{26}$$

8.4 Initialization

The initialization is the same as that described above, although the initial conditions matrix may become more stacked than before when other items are added, except that here $\beta_1 \sim NID(0, \kappa),\ where\ \kappa \to \infty$ also.

8.5 Adding regressors and event indicators

Not only are these equations stacked as such, the state vector can be a stack of unobserved component, regressor variables, and event dummies. This state vector of components can now be augmented by regressors and dummy indicators.

At each time period, the state vector, α_t, is projected forward in time by the two equations preceding it. This defines the temporal evolution of a dynamic factor model.

$$\alpha_{t+1} = T\alpha_t + w_t b + RG_t \tag{27}$$

where w_t = the regression parameters, b= the regression coefficient matrix, T = the transition matrix, R = selection matrix of zeros and ones. G_t = an error vector, and α is a state vector of stacked components.

Our initial measurement equation may consist of one or more components (local level and local slope), plus explanatory variables (perceived Chornobyl health risk and recalled count of illnesses), and potentially some event dummy variables (identifying outliers or level shifts):

$$y_t = \mu_t + \beta_t + \lambda_t I_t + \omega x_t + H_t \quad H_t \sim NID(0, \sigma^2_{H_t}) \tag{28}$$

where μ_t = local level component; β_t = slope vector component (gradient of trend, μ_t); ω_t = vector (of explanatory mean-centered variables, x_t); λ_t = vector of coefficients (of event indicators, I_t); and H_t = measurement error vector. In this case this model, there would be multiple stacking of components, fixed and/or stochastic variables, and dummy impact indicators on top of one another in the state vector.

8.6 Augmented Kalman filter

Further stacking of the non-stationary and stationary components takes place within the state vector, so that a diffuse prior can be applied to the partition of non-stationary components while normal maximum likelihood procedures are applied to the stationary partition [13], [14]. Readers interested in the details of this recursive algorithm are encouraged to read excellent discussions of Harvey(1989) [22], ,Commandeur and Koopman(2007) [11], and Durbin and Koopman(2001) for background [17].

8.7 Analytical protocol

We analyzed our models in six stages. In this initial stage, we performed our hypothesis testing. In the first stage, we attempted to fit an local linear trend model along with all of the explanatory variables we attempt to test. We tested the average cumulative external dose in mSv, the perceived risk score, and the mean self-reported number of illness per period. We retained all of the components with significant signal to noise ratios. If the model converged, we retained all of the statistically significant components and variables.

In the second stage, we refined our model by trimming out the superfluous components. We removed the variables that did not contribute to a proper fitting of the model and estimated the trimmed model with maximum likelihood until the model fully convergence to a steady state. We endeavored to refine the model by adding event dummy effects that significantly improve the fit. By testing the significance of temporal events with intervention variables, we determined what events with immediate impact significantly explained the endogenous depression variable, y_t . In the trimming of the model, we refined it so as to maximize the variance explained by the endogenous variable.

Thirdly, we assessed the model for validity by subjecting it to a series of misspecification tests. If the model passed these tests, the assumptions were adequately fulfilled so the model could be considered valid. These tests confirmed the independence, homogeneity, and normality of the residuals of the components estimated.

Fourth, we evaluated we the last eight observations of the model for predictive validation. In these tests, we tested for significant deviation of the forecast over the observed data within the validation segment of the sample. If the model passed these tests, we deemed it suitable for further forecasting.

Fifth, reverting to the 2005 point of forecast origin, we proceeded to forecast over the remaining five years to circumvent the surges in depression that could confound or analysis.

Sixth, we subjected the final conditional forecasts to forecast evaluation with one novel modification. We began our forecasting prior to end-effects that could confound our analysis. We evaluated the accuracy of the conditional forecasts against the realized depression and show that the end-effects exceed 95% forecast confidence intervals from a 2005 point of forecast origin. We assumed that these surges in depression follow from a cascade of new crises. Although we do not claim that this method circumvents all intervening variables, we employ it to circumvent intervening impacts of crises that could undermine the internal validity of a retrospective analysis. We will discuss this assumption in detail later.

9 Sample Characteristics

The sample consisted of 702 respondents from the Kiev and Zhitomyr oblasts in Ukraine. By gender, the sample comprised 48.29 % (339) males and 51.71% (363) females. Most of the sample resided in Kiev oblast at the time of the survey (2009 through 2011) 85.9% (603) came from Kiev and 14.1% (99) from Zhitomyr oblasts. Approximately 32%(224) were 41 years old or younger. Approximately 31% (218) were 42 thru 54 years old, and about 37%(260) of the sample were 55 or more years old.

At the time of the interview, most respondents (approximately 70%(488)) reported being married with 9% (64) reporting being single. About 7% (49) said that they were divorced and about 9% (61) reported being widowed. About 5%(32) indicated that they were cohabiting.

The occupational status consisted of a plurality percentage as occupying a professional, executive, or administrative position at the time of the interview. Approximately 62% (434) indicated that they were employed full time with about 8.12% reporting part time employment. Only 5.13% claimed that they were unemployed and about 25% described themselves as being retired. At the time of the interview about 26.92% (189) of the respondents indicated that they were professional, executive, or administrative positions. 26.63% (167) reported that they were homemakers or caregivers. 19.30% (121) described their positions as technical sales or administrative support. 11.96% (75) maintained that they were in a service occupation or protective services. Almost 7% (43) stated that they were working in a precision production or mechanical craft or construction. 3.67% reported being factor laborers, machinists, cleaners, or involved in transportation. 1.28% (8) indicated that they were involved in agriculture, forestry, fishing, trapping, or logging. One respondent reported being a student.

The sample was in general well educated, with slightly more than 40% (281) indicating that they had finished a master's or specialist's degree. 34.33% (241) reported having a technical degree. 13.82% (92) indicated that they had graduated college with a bachelor's degree. About 1% indicated that they held a PhD, Ph.sci, or MD. The drop-out or incomplete rate was minuscule in that only 5.27% (37) reported having only a high school or less education and only 5.41% had not finished college or had not earned a bachelor's degree. In terms of education, this was a fairly well-accomplished sample.

Income sufficiency indicated the financial stress on the sample. Approximately 15.78% (98) of the responses indicated that their income was not sufficient for basic necessities, whereas 47.67% (296) reported of the responses reported that their income was just sufficient for the basic necessities. About 1/3 (33.01%)of the responses (205) maintained that their income was sufficient for the basics plus a few extra purchases or savings now. Only 3.54%(22) of the responses admitted that they live comfortably and afford luxury items.

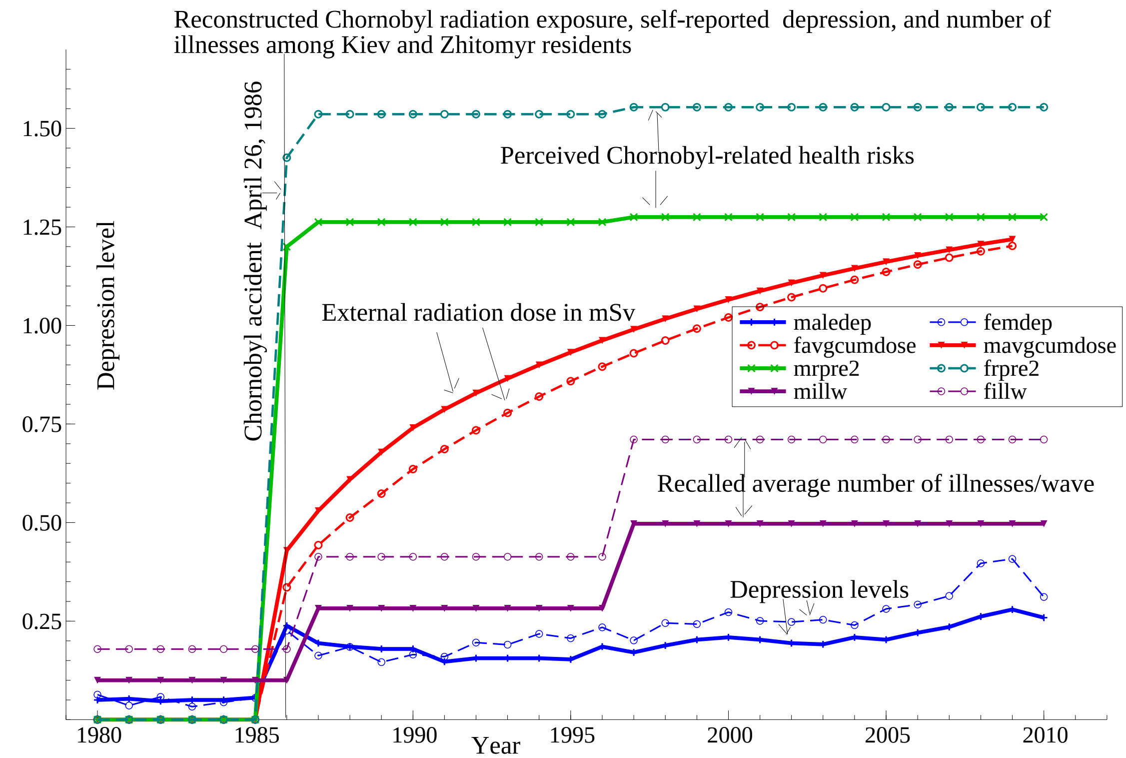

Figure 2: Cumulative external dose (excluding natural ambient and medical diagnostic or treatment radiation) and psycho-social depression in Kiev and Zhitomyr oblasts

10 Psycho-social depression in Kiev and Zhitomyr oblasts following the Chornobyl nuclear accident

We examined psycho-social depression of the residents of Kiev and Zhitomyr oblasts separately for males and females for the effect of Chornobyl on that symptom. We compared the impact of Chornobyl and we forecast from immediately after the Orange revolution at the end of 2004, using 2005 as the point of forecast origin. In this way, we attempted to control for the impact of events after 2005, which are associated with a surge in depression. Without such a control, the confounding impact could render any semblance of the Chornobyl effect hopelessly distorted. In Figure **??**we displayed the levels of depression before, during, and after Chornobyl with a view toward identifying immediate impacts on these events.

It should be noted that the external dose not only does not include natural background radiation (from radon in the ground or cosmic radiation), nor does it include any radiation from voluntary medical or dental x-rays either.

10.1 Male Depression Model

For the inquiry into male depression, we began an analysis of an initial model, consisting of the unobserved components and the regression predictor variables– namely, average cumulative external dose measured in mSv, perceived Chornobyl health risk,

The conventional formula applies only to stationary data, where values are bounded by the limits of 0,1. For non-stationary data, Andrew Harvey (1989) suggested using a variant that is more robust to trending data, which he calls Rd^2:

$$Rd^2 = 1 - \frac{sse}{\sum_{t=1}^{T}(\Delta y - \Delta\bar{y})^2} \tag{29}$$

where the denominators consist of differenced terms measuring rates rather than undifferenced ones measuring levels of the endogenous variable [22, 268].

The Rd^2 is a formula for proportion of explained variance, where the denominator of the proportion of error variance is more robust to nonstationary data than the denominator in the conventional R^2 formula.

Both the omnibus fit statistics, the State vector components, and the State vector regression effects of the initial and final models are displayed in Table 3.

The trimmed male depression model included the same unobserved components in our final model, along with the parameters that were found to be statistically significant, plus any event dummies that help explain significant variations in the residuals. Non-significant parameters are usually pruned from this model. In the trimmed model, we changed one intervention, leaving a more parsimonious explanation of depression.

10.1.1 Initial male model depression model

Our initial male depression model contained the a local level, local slope, and an irregular component, plus some explanatory variables and event impact indicators. This model tested the effects of male average age, cumulative external dose effects, male perceived Chornobyl related health risk and recalled average number of illnesses per period, along with an event impact dummy variable for a level break in 1991 contained within it.

The general form of the trimmed male depression model is only slightly different in that it lacks a significant slope component. The trimmed male depression model fit about as well as the full model, as shown in Table 3, but is much more parsimonious. We pruned statistically non-significant variables to obtain a more parsimonious model. In addition to dropping the average cumulative male dose, we added an event impact dummy at 1996 to significantly improve the fit. The parameter estimates of both models are displayed in Table 3.

We found that the average male cumulative external dose of radiation from ^{137}Cs did not appear to have a statistically significant impact on male depression in this sector of Ukraine. After it appeared to be non-significant in the initial model, we pruned it from equation. Male psycho-social depression dropped in 1991 as the Soviet Union collapsed. However, perceived Chornobyl related health risk remained as statistically significant. The coefficient is positive, indicating that there appears to have been a positive relationship between perceived Chornobyl related health risk for men and male psycho-social depression in Kiev and Zhitomyr oblasts. There was a positive shift in depression in 1996.

Thus the measurement equation formula that appears to parsimoniously explain and predict male depression can be expressed as

Table 3: **Male Depression Models with Regression effects in final states**

	Initial	**Full**	**Male**	**Model**	**Final**	**Trimmed**	**Male**	**Model**
BIC	-8.104				-8.444			
LL	85.354				- 90.979			
-2LL	-170.707				-181.958			
Rd^2	0.956				0.953			
Prediction Error variance	9.7e-05				8.95e-5			
State vector components	**value**			**prob**	**value**			**prob**
Level	0.274			0.254	0.094			0.001
Slope	0.003			0.003	0.002			0.601
Final State regression effects:	**Coef**	**SE**	**t-value**	**prob**	**Coef**	**SE**	**t-value**	**prob**
Mean cumulative external dose	-0.133	0.152	- 0.860	0.402				
Recalled # of Illnesses	-0.116	0.051	-2.330	0.033	-0.177	0.037	-4.769	0.000
Male perceived risk	0.196	0.0512	3.784	0.001	0.152	0.009	17.662	0.000
dlmaleage	-2.041	9.875	-0.207	0.839				
Level shift 1987	-0.034	0.014	-2.554	0.020				
Level shift 1991	-0.036	0.011	-3.150	0.006	0.033	0.010	3.180	0.005
Level shift 1996					0.028	0.010	2.748	0.012
Final Status	Steady	state	w full	converge.	Steady	state	w full	converg.

Table 4: **Misspecification tests**

Misspecification Test	**Sig level**
residual autocorrelation ACF of standardized residuals	ns
Bowman -Shenton test residual normality	ns
residual homogeneity cusum residuals	ns
Irregular residual normality	ns
Irregular residual skewness	ns
Irregular residual kurtosis	ns
Level residual normality	ns
Level residual skewness	ns
Level residual kurtosis	ns
Slope residual normality	ns
Slope residual skewness	ns
Slope residual kurtosis	ns
Legend ns = ($p > 0.05$)	**if** $p < .05$**, then p-value is displayed**

$$\begin{aligned} MaleDep_t &= 0.094Level_t + 0.002Slope_t + Irregular_t \\ &\quad -0.177NumIll_t + 0.152MalePerceivedRisk_t \\ &\quad +.033LevelShift\_1991_t + 0.028LevelShift\_1996_t \end{aligned} \tag{30}$$

where $MaleDep_t$ = male self-reported depression above 5%, $Level_t$ = level component, $Slope_t$ = nonsignificant slope component, $NumIll_t$ = average number of recalled illnesses per wave, $MalePerceivedrisk_t$ = Male Chornobyl related perceived health risk, and $LevelShift\_1991_t$ = level shift at collapse of USSR, and $Levelshift\_1996_t$ = level shift at completion of Constitution and creation of national currency. Level shifts are event indicator dummy variables coded as 0 before the event, and 1 thereafter.

From this equation for male depression, it is clear that the dominant explanatory variable is that of male perceived health risk relating to exposure to radioactivity from radiation released during the Chornobyl accident. Cumulative external dose is not a significant explanatory variable for the men. We notice that the recalled average number of illnesses per wave emerges as a statistically significant explanatory variable with a negative coefficient. Perhaps the identification or diagnosis of physical illness allows treatment and recovery, reducing reason.

10.1.2 Male depression model validation

To test for misspecification, we tested for residual autocorrelation with standardized residuals, for residual normality with a Bowman-Shenton test, along with tests for kurtosis and skewness. All tests were passed as summarized in Table 4.

Figure 3: Male depression after Chornobyl

When we find that this model fits well, we examine the estimated signal when graphed against the data, in Figure 3.

10.1.3 Male depression model validation

Although the model fit quite well, we had to be sure that it is not overfit if we were to forecast with it. We performed two tests for predictive validation. We forecast over the last 8 years of data and compared the data to our forecasts to obtain an *ex post* forecast evaluation. We displayed the results of these two tests in Table 5.

The results of the two statistical tests are encouraging. We found no statistically significant difference between the forecasts and the data with a χ^2 (df=8) test. Nor did we find that the forecast exceeded the bounds of significance in the CUSUM test. Having reason to believe that our forecasts were reasonably accurate, we performed our final forecast from 2005 onwards.

10.1.4 Some Ukrainian historical highlights

Before evaluating the forecast, we noted some of the Ukrainian historical highlights between the accident at Chornobyl and patterns of general psycho-social male depression after 2005[4] that accompanied these events. We used the historical highlights as temporal markers rather than as imputed causes.

We noticed a large spike in male depression in 1986, the same year as the accident at Chornobyl. We found no delay in the rise of male depression at the date of the Chornobyl accident. Although the level of male depression dropped substantially during the following year. For the next two years, the rate of male

Table 5: **Male depression model Predictive validation**

Predictive validation					
year	**error**	**stand.err**	**residual**	**cusum**	**sqrsum**
1998	0.014	0.011	1.252	1.252	1.569
1999	0.008	0.011	0.691	1.194	2.046
2000	-0.003	0.011	- 0.269	1.675	2.118
2001	-0.014	0.011	-1.294	0.381	3.793
2002	-0.014	0.011	-1.301	- 0.921	5.487
2003	-0.005	0.011	-0.497	- 1.418	5.734
2004	0.016	0.011	1.506	0.088	8.001
2005	-0.010	0.011	-0.968	-0.881	8.939
Ex post-sample predictive evaluation.					
Chi2(8) test is	8.934	[0.346]			
Cusum t(8) test is	-0..311	[1.237]			
Ex post-sample prediction statistics.					
Sum of 8 absolute prediction errors	0.085				
Sum of 8 squared prediction errors is	0.001				
Sum of 8 absolute prediction resids	7.779				
Sum of 8 squared prediction resids	8.939				

depression decline lessoned, and there was a steep decline in 1990, the year before the collapse of the Soviet Union. In 1991, we observed a slight increase in male depression which flattened out till 1993 and lessens in 1994, the year of presidential elections.

In 1994 Leonid Kuchma succeeded Leonid Kravchuk as President of Ukraine. Depression increased. In 1996, a new democratic constitution was adopted and a new currency was introduced.

In 1997 a friendship treaty with Russia was concluded along with an agreement about a port for the Russian Black Sea fleet. In 2000 the Chornobyl nuclear plant was finally shut down, and in 2002, the Ukrainian leadership announced a bid to join NATO.

In 2004 Yanukovich won an election that observers claimed was rigged and which the supreme court later annulled. Victor Yushchenko won an election and took office in 2005. Yulia Tymoshenko became prime minister, but by September 2005, President Yushchenko dismissed her government. We notice that average male depression increased after 2005, shown in Figure 4.

10.1.5 A Period of Pronounced Ukrainian political and economic turmoil

The period of 2006 through 2009 was riven with political and economic crisis in Ukraine. A Russian-Ukrainian natural gas dispute developed during 2005, which led to a January 2006 Russian cut-off of the supply of natural gas.

From 2006 to 2007, we observed a decline in average male depression. In 2007 we discerned average male depression increase, in Figure 4, the data for which is provided in Table 6.

In 2006, a new constitution and currency(Hrivnia) was introduced. However, the winter was characterized by an arctic blast that killed 53 people. Moreover, a hepatitus outbreak took place in the Crimean part of

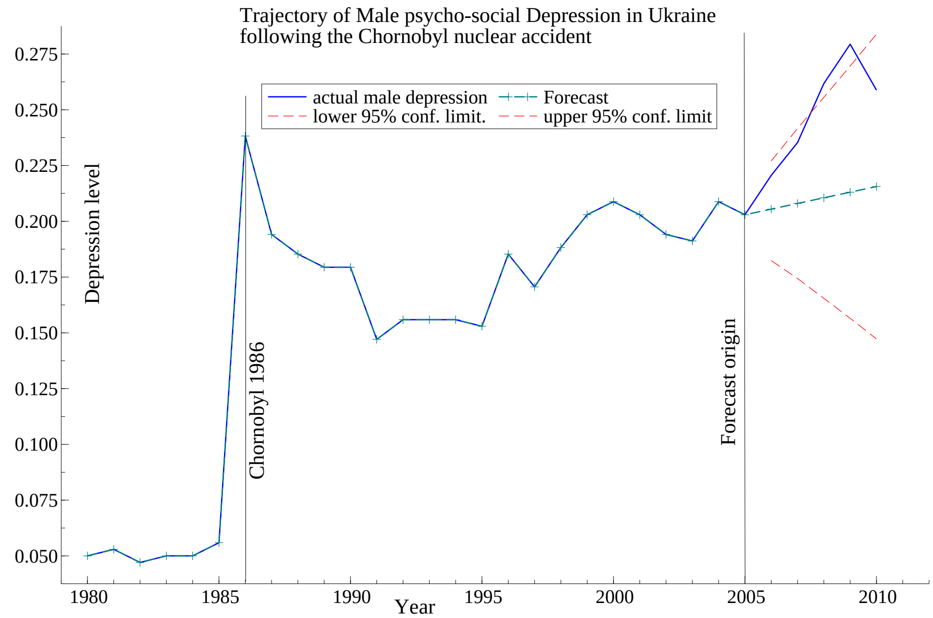

Figure 4: Ukrainian male depression forecast profile with 95% prediction confidence limits in Kiev and Zhitomyr oblasts 2005-2010

Ukraine which hospitalized approximately 3000 people.

This year of 2007 was a period of political crisis and struggle in Ukraine. President Yushchenko tried to dissolve the parliament and call for elections in May. Parliament called this decree unconstitutional and refused to fund the new elections. Before the Court could rule on the issue, Yushchenko dismissed three of the Constitutional Court judges. Yushchenko then cancelled his previous decree and rescheduled the elections for September. In September one of the judges was reinstated by the court and the funding for an election was arranged.

The year of 2008 was one when the impact of the Great Recession was experienced. In 2008, Gazprom agreed to a new contract to supply Ukraine with natural gas in March, but by October 2008, the global financial crisis led to a decline in steel prices which depressed the demand for Ukraine's main export. Unemployment rose and the Ukrainian currency value fell further.

By January 2009, the Russians cut off the gas again, which forced the closure of much of the industrial sector of the Ukraine. Eventually in 2009 Tymoschenko brokered a gas treaty with Russia.

Instead of allowing these multiple economic and political crises of 2006 through 2009 to significantly inflate reported depression, we forecast from 2005 onward, obtaining 95% forecast confidence limits as upper boundaries of psycho-social depression before this new era of aggravation of depression. As noted above, there were a multitude of possible causes of this increasing depression. We did not try to identify all of them or even control for sources of minor variation. We merely endeavored to avoid any confounding of reported depression from the impacts of historically intervening multiple, major crises.

Table 6: **Male Depression Forecast profile from 2005 onward**

	Forecasts	**with**	**95% confidence interval**	
period	**Forecast**	**stand.err**	**leftbound**	**rightbound**
2006	0.20547	0.01079	0.18389	0.22705
2007	0.20800	0.01682	0.17436	0.24164
2008	0.21053	0.02253	0.16547	0.25559
2009	0.21306	0.02827	0.15652	0.26960
2010	0.21559	0.03416	0.14727	0.28391

10.1.6 Forecast evaluation of male depression model

We needed to evaluate this forecasting model for its accuracy over the *ex ante* forecast horizon. Because our temporal reversion to an earlier point of forecast origin, we could evaluate our *ex ante* forecast against actual data. We already knew that this was a time of multiple political and economic crises associated with inflated the actual reported depression, so we are not worried about a less than perfect forecast evaluation. We expected the end-effect in the forecast horizon to enhance the error and this was the motivation for circumvention of this potential distortion of long-lasting reported depression. Nonetheless, we have to obtain a sense of how well our forecasts work under such circumstances.

To evaluate our final forecasts, we used several criteria of forecast accuracy. We considered the percent coverage of the real data by the 95% forecast intervals, the root mean square error (RMSE), the root mean square prediction error (RMSPE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE), listed in Table 7. These criteria enabled us to compare the results in our discussion at the end of the paper. Instead of concentrating on forecast accuracy, we use the MAE and the MAPE, under these circumstances, to assess, respectively, the absolute and relative inaccuracy of our forecasts, and by extension, the absolute and relative impacts of the the multiple crises, on reported depression.

The formula for the mean absolute error is

$$MAE = \frac{1}{h} \frac{\sum_{t=1}^{h} |F_t - A_t|}{\sum_{t=1}^{h} |A_t|} \tag{31}$$

and the formula for the mean absolute percentage error is

$$MAPE = \frac{100}{h} \frac{\sum_{t=1}^{h} |F_t - A_t|}{\sum_{t=1}^{h} |A_t|} \tag{32}$$

where F_t = forecast for time period t, A_t = actual value of observation at time period t, and h = number of time periods in forecast horizon.

Although the MAPE puts different scales on a uniform scale of percent, it suffers from scale dependence. When the MAPE applies to forecasts of very small numbers– such as errors of size 1 or 2, they become converted to large percentages. For example, if the correct value is 1, but the forecast is 2, an error of 1 unit is 100% in error. Similarly, if we found that the MAE was less than five, we could make an explicit allowance for an apparently large percentage error due to scale dependence of the MAPE.

To compensate for such scale dependence scholars have proposed a symmetric MAPE (SMAPE). One version (number 3) of this SMAPE is more robust to outliers than others. The formula for the SMAPEv3 refers to the the portion of the data in the forecast horizon.

Table 7: **Male Depression Ex Ante Forecast Evaluation**

Forecast accuracy	measures	from	2005	forwards:	
period	**Error**	**RMSE**	**RMSPE**	**MAE**	**MAPE**
2006	-0.01512	0.01512	0.68531	0.01512	6.85314
2007	-0.02729	0.02206	0.95267	0.02121	9.22639
2008	-0.05123	0.03463	1.37186	0.03121	12.67511
2009	-0.06635	0.04472	1.67966	0.04000	15.44302
2010	-0.04323	0.04443	1.67781	0.04065	15.69516

$$SMAPEv3 = \frac{100 * \sum_{t=h-1}^{H} |F_t - A_t|}{\sum_{t=h-1}^{H} |F_t + A_t|} \tag{33}$$

where H = the total number of periods in the forecast horizon and h=the index for period number of the forecast horizon. When we apply this to the forecast horizon and the period immediately before it, we observe that the SMAPE for males is only 2.96%, a smaller and more tolerable error than that found in the maximum estimate of the MAPE. With this much error in the *ex ante*forecast, it appears that the accuracy of this projection is much better than we would have thought, had we relied on the MAPE alone for forecast evaluation.

10.2 Female Depression Model

10.2.1 Initial female depression model

The initial female model included a local level, a slope, and an irregular component. Neither the local level nor the slope component was significant in this initial model, even though the signal appeared to track the data quite well in Figure 5.

The parameter estimate details of the initial and final female depression model can be found in Table 8.

In this initial full model, we tested the reconstructed external female dose (Mean cumulative external dose, and found this effect to be statistically non-significant. Nor did was the first difference of the natural log of female age (`dlfemage`) to be significant . The only variable found to be statistically significant in the initial full female model was female perceived related health risk of exposure to radiation from the radioactivity released during the Chornobyl accident.

We also tested the slope parameter only to find that this was not statistically significant either. Both of these terms became candidates for deletion in the final model. The slope parameter was not found to be a statistically significant component in the trimmed model.

To improve the fit of the initial female model, we trimmed the nonsignificant external dose variable and the female self-report of average number of illnesses per wave. The variation accounted for by it was apparently explained by the local level. What remained in the model was the local level, along with the other variables we were testing–including the female mean cumulative external dose, the recalled number of illnesses per wave, and the female perceived risk from exposure to radioactivity from the radiation released during the Chornobyl accident, plus a level shift variable at 1987.

The trimmed female depression model is a more parsimonious explanation of the phenomena under consideration and it fit better than the initial model. What was left is a local level and (almost significant)

Table 8: **Female Depression Models with Regression effects**

	Initial	**Full**	**Model**		**Final**	**Trimmed**	**Model**	
BIC	-6.879				-7.390			
LL	76.501				77.535			
-2LL	-153.002				-155.069			
Rd^2	0.845				0.886			
Prediction error variance	0.0004				0.003			
State vector components	**value**			**prob**	**value**			**prob**
Level	0.003			0.993	0.165			0.000
Slope	0.016			0.268	0.005			0.096
State regression effects	**Coef**	**SE**	**t-value**	**prob**	**Coef**	**SE**	**t-value**	**prob**
Mean cumulative external dose	-0.136	.211	-0..644	0.528				
Recalled # of Illnesses	-0.088	0.067	1.309	0.207				
Female perceived risk	0.152	0.047	3.254	0.004	0.066	0.013	5.237	0.000
dlfemage	12.800	15.060	0.850	0.407				
Level Shift 1987	-0.048	0.0318	-1.515	0.147				
Outlier 1986					0.075	0.018	4.134	0.001
Outlier 1997					-0.031	0.015	-1.997	0.060
Outlier 2000					0.028	0.015	1.813	0.085
Final Status	Steady	state	w full	converg.	Steady	state	w full	converg.

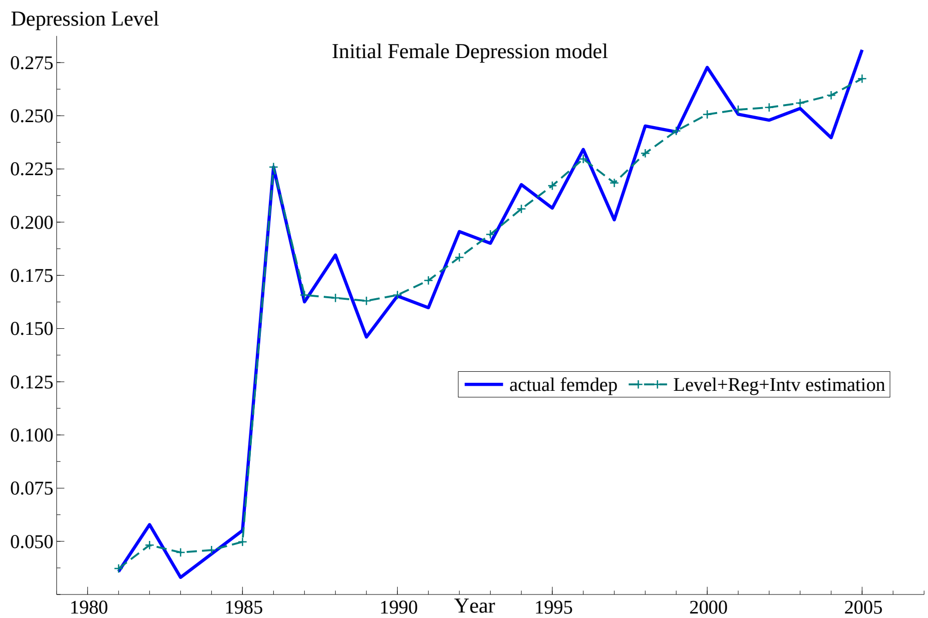

Figure 5: Female depression after Chornobyl

stochastic slope model. We also retained the female perceived risk of exposure to radiation released during Chornobyl and the blip dummy representing a spike in female depression during shock of the Chornobyl accident of 1986. However, we added another event dummy at 1996 to significantly improve the fit. The resulting model parameters are displayed in Table 8.

10.2.2 Trimmed female depression model

Our trimmed female depression model included the local level, the female perceived risk, and three outliers– one in 1986 at the time of Chornobyl, another in 1997 and another in 2000.

The final female measurement depression model was formulated as

$$\begin{aligned} FemDep_t = 0.165 Level_t + 0.005 Slope_t + Irregular_t \\ +.067 FemPerceivedRisk_t \\ +0.075 blip1986_t - 0.031 blip1997_t + 0.028 blip2000_t \end{aligned} \tag{34}$$

where Fem Dep_t = Female reported depression, $Level_t$ = local level component, $Slope_t$ = statistically almost significant local slope component, $Irregular_t$= irregular (error or noise) component, $FemPerceivedRisk_t$ = female perceived Chornobyl related health risk. and $blip1986_t$ is a blip dummy indicator, coded as 0 when the event is not taking place and 1 during the occurrence of the represented event. In this equation, the outlier indicators are coded as 1 for the year indicated and zero otherwise.

When we examine the trimmed female model, we observe that the largest effect appears to be that of the time-varying level. The second largest coefficient is that of the 1986 spike in depression ($blip\_1986_t$) at

Table 9: **Female model misspecification test results**

Misspecification Test	**Sig level**
residual autocorrelation ACF of standardized residuals	lag3 : 0.01
Bowman -Shenton test residual normality	ns
residual homogeneity cusum residuals	ns
Irregular residual skewness	ns
Irregular residual normality	ns
Irregular residual kurtosis	ns
Level residual normality	ns
Level residual skewness	ns
Level residual kurtosis	ns
Slope residual normality	ns
Slope residual skewness	ns
Slope residual kurtosis	ns
Legend ns $= (p > 0.05)$	**if** $p < .05$**, then p-value is given**

the time of the Chornobyl accident. Almost as great as this impact is that of the female perceived risk of exposure to radioactivity from radiation released from the Chornobyl event. It is noteworthy that female cumulative external dose is not a significant predictor of reported female depression in this model. The negative coefficient of the 1997 outlier may express establishment of the democratic Constitution in 1996, the establishment of the national currency, and the signing of the Ukrainian-Russian friendship pact in 1997.

10.2.3 Female depression model validation

To assure proper specification, we tested for residual autocorrelation with standardized residuals, for residual normality with a Bowman-Shenton test, along with tests for kurtosis and skewness. All tests were passed with the exception of some significant outliers in the irregular residuals and the results are displayed in Table 9.

10.2.4 Female depression model predictive validation

The model appeared to be generally well specified, apart from a significant autoregressive lag 3. When we assessed the trimmed model for predictive validity before further forecasting, we obtained the results tabulated in Table 10. That significant third lag appears to have had not to have significantly affected the predictive validity of the model.

10.2.5 Forecast profile for Ukrainian female depression

Before forecasting with our data, we performed an *ex-post* forecast evaluation, the test results for which are contained inTable 10 . We found that there is no significant difference between the forecasts and the data,

Table 10: **Female depression predictive validation**

Predictive validation					
year	**error**	**stand.err**	**residual**	**cusum**	**sqrsum**
1998	0.008	0.021	0.386	0.386	0.149
1999	-0.006	0.019	-0.337	0.049	0.263
2000	0.000	.	0.000	0.050	0.265
2001	-0.007	0.021	-0.342	-0.292	0.385
2002	-0.011	0.019	-0.592	-0.882	0.738
2003	-0.005	0.019	-0.262	-1.146	0.808
2004	-0.021	0.018	-1.147	-2.294	2.115
2005	0.027	0.018	1.483	-0.8104	4.315
Ex post-sample predictive evaluation.					
Failure Chi2(7) test is	4.315	[0.743]			
Cusum t(7) test is	-0.310	[1.232]			
Ex post-sample statistics.					
Sum of 7 absolute prediction errors is	0.086				
Sum of 7 squared prediction errors is	0.001				
Sum of 7 absolute prediction resids is	4.550				
Sum of 7 squared prediction resids is	4.315				
Legend:	1 missing value = .				

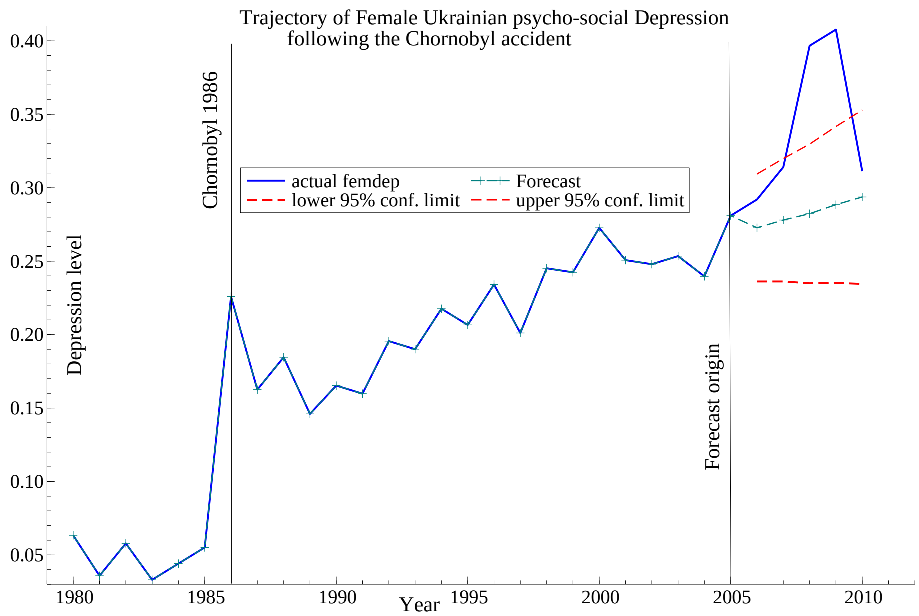

Figure 6: Female depression Forecast profile from 2005 over a five year horizon

and that the forecasts did not exceed the boundaries of the cusum test. With confidence in our predictive validation, we proceeded to forecast ex ante.

We needed to evaluate this forecasting model for its accuracy over the *ex ante* forecast horizon, which, because of our temporal reversion to an earlier point of forecast origin, permitted us to evaluate an *ex ante* forecast.

We observed that the depression level emanating from Chornobyl and events shortly thereafter can be projected ahead up to the current time. From this forecast we observed that the actual depression level exceeded that forecast up to 2013.

Whether the reason for this excess depression is political or economic or sociological or psychological clearly the impact of events after 2005 is contributing to excess depression on the part of the female Ukrainian depression. One thing was probably the Great global recession in September 2008. It was all downhill after that. So depression naturally increased. When the gas was cut-off on January 1, 2009, depression reached a global maximum. We needed to examine the forecast profile in Table 11.

10.2.6 Forecast evaluation of female depression model

To evaluate the forecasts we use the error, the root mean square error, the root mean square prediction error, the mean absolute error, and the mean absolute percentage error, listed in Table 12. Moreover, we can use these measures to compare the forecasts between males and females over time.

However, the SMAPEv3 for the female depression forecast is only that of 4.73%, which indicates more forecast accuracy than would have been suggested the MAPEe alone.

Table 11: **Female Depression Forecast profile from 2005 onward**

Forecasts	**with 95%**	**confidence interval**	**from**	**2005 forwards:**
period	**Forecast**	**stand.err**	**leftbound**	**rightbound**
2006	0.27277	0.01829	0.25450	0.29109
2007	0.27801	0.02092	0.25711	0.29898
2008	0.28324	0.02369	0.25956	0.30702
2009	0.28847	0.02660	0.26189	0.31519
2010	0.29371	0.02962	0.26409	0.32348

Table 12: **Female Depression Ex Ante Forecast Evaluation**

Forecast accuracy	measures	from	2005	forwards:	
period	Error	RMSE	RMSPE	MAE	MAPE
2006	-0.01924	0.01924	0.65888	0.01924	6.58882
2007	-0.03604	0.02889	0.93578	0.02764	9.03300
2008	-0.11345	0.06962	1.81942	0.05625	15.55529
2009	-0.11914	0.08479	2.14965	0.07199	18.97790
2010	-0.01759	0.07625	1.93923	0.06111	16.31214

11 Discussion

11.1 Modeling Trajectories of Depression

Our random telephone sample of the Ukrainian residents of Kiev and Zhitomyr oblasts focuses on psycho-socio medical findings characteristic of the Ukrainian public. The sample did not focus specifically on cleanup workers, evacuees from the exclusionary zone surrounding Chornobyl nuclear power plant, or young children whose thyroid uptake of ^{131}Iodine may have led to excessive thyroid problems among this population segment. In Figures 1 and 2, we displayed the paths of the other variables that we tested to determine whether they explained and predicted the reported levels of depression.

We graphed the trajectories of the reported levels of depression for the respondents from 1980 to 2010 in Figures four and six. For both males and females, these figures exhibit two principle surges in the level of depression. The first surge occurred in 1986 during and immediately following the Chornobyl accident. The second surge appeared between 2005 and 2010. After the first surge, the average level of depression never returned to the to the pre-accident level. For both males and females, we found that the exposure to radioactivity released during the Chornobyl accident was not a significant variable in a model either in the explanation or in the prediction of this psycho-social depression. Rather, the perceived health risk from exposure to radiation from radioactivity released during that accident is a dominant explanatory predictor variable in both models explaining depression trajectories over time.

This empirical finding is supported by a report from the US National Academies of Sciences: Biological Effects of Ionizing Radiation(BEIIR VII)(71) which summarizes the latest findings of epidemiological and experimental research on low levels of ionizing radiation. At doses less than 100 mSv, statistical limitations make it difficult to evaluate risks in humans. The lifetime-attributable-risk (LAR) for developing cancer in a population receiving 100mSv would be about 1% for males and 1.4% for females. This is about 40 times lower than the incidence of cancer expected in the population from other causes. For our representative sample, the median accumulated dose is about 10 times less than the external dose expected from naturally occurring background sources.

We compared the male to the female depression trajectories in Figures 4 and 6. We found that both surged to new peaks at the time of Chornobyl and declined sharply in the following year. Both paths irregularly declined until 1991.

By 1991, the male and female depression paths crossed, as the male path dropped below that of the female. By 1991 both depression trajectories began an irregular climb upward. The female path alternately rose and fell twice, but each time it rose, it moved twice as far as it fell. In 1991, the male depression path rose a little, remained level for two years before falling slightly. In 1995, both paths rose for a year and then fell the next year. Both paths reached a local maximum in 1996, (after the arctic chill killed 53 people and the hepatitus outbreak in Ukrainian Crimea hospitalized about 3000) and then declined somewhat in 1997 before both trajectories began to rise again. They attained a new local maximum in 2000 after which they began declining until they reach a new turning point around years 2004-2005.

From 2005 onward, both depression trajectories rose until 2007 when they begin to surge upward again. By 2008-2009 the depression levels reached a new peak, greater in magnitude than the depression level reached in 1986.

With these data, we have developed state space models, based on a local level component, on the perceived risk from radiation exposure to radioactivity released during the Chornobyl accident, plus some event intervention dummy variables, formulated in Equations 30 and 34, that can track and forecast these psycho-social depression trajectories following the nuclear accident at Chornobyl. Thus, we developed state space models of psycho-social depression on the part of Ukrainians after the Chornobyl accident, and we used these to test whether actual exposure or perceived risk predicted such depression, finding that the latter was the primary driver of these depression. Moreover, we demonstrated that we could revert to an earlier point of forecast origin to circumvent the major impacts of mediating events at the end of our time series that otherwise would have impaired the internal validity of our analysis.

11.2 Predictive validation

We assessed the male and female model of reported depression for predictive validation. We examined the validation period of six to eight years before the point of forecast origin, and we found that our forecasts over this validation period were not significantly different from our actual data, which provided justification for further forecasting.

11.3 Choice of a point of forecast origin

The period of 2006 through 2010 was an era of multiple major political and economic crisis in Ukraine. As depicted in Figures 4 and 6, a surge in depression levels occurred from 2006 through 2009. We saw the impact of these confounding variables as we observed the rise in realized depression that exceeded the upper 95% confident limit of the depression forecast from the point of origin, 2005, to the end of our data, in 2010. This surge in reported depression is especially pronounced for women from 2007 until 2009. In order to avoid confounding these other sources of disconcertment with depression, we decided to select our point of forecast origin early enough in time to avoid having to model events after 2005.

Table 13: **Male depression forecast evaluations from 3 different points of origin**

Forecast origin year	Horizon length	Percent coverage	MSFE	RMSFE	MAE	MAPE
1991	19	0.0000	0.134	0.366	0.346	166.863
1996	15	92.857	0.001	0.026	0.019	7.848
2005	5	60.000	0.002	0.044	0.041	15.700

11.4 Forecast accuracy over 2005-2010 forecast horizon

For our forecasts to be considered reliable, we have to assess them for accuracy. We do this by evaluating the forecasts according to three measures of forecast accuracy over the 2005 to 2010 forecast horizon. They are the percent coverage of the real data by the 95% confidence intervals, the MAE, and the MAPE. We use the absolute value criteria rather than mean square measures because the former do not exaggerate errors by squaring them.

When we examined our forecasts over the 2005 to 2010 horizon, we found the forecasts to be reasonably accurate for both men and women, given that the surge in actual depression exceeded the 95% upper forecast limit. In Tables Tables 13 and 14 forecast accuracy measures for males and females respectively over that five year horizon are listed. The MAE and MAPE for male respondents are 0.041 and 15.7, respectively, whereas for the women they are 0.0621 and 16.312. The percent of non-coverage, shown as the partial exclusion of the actual depression data by the 95% forecast intervals, in Figures 4 and 6, was estimated to be 40 percent for both males and females. Most importantly, the SMAPES for males and females were respectively 2.96% and 45.73%, both less than 5.00 %, So the use of temporal reversion to a more optimal point of forecast origin does not appear to have impaired this second horizon forecasting.

11.5 Alternative points of forecast origin

We tested the accuracy of the of the forecasts from 1991, when the Soviet Union collapsed, and from 1996, when the Ukrainian constitution was written and the national currency established, against those from 2005, and summarized the forecast accuracies in Tables 13 and 14.

The use of 1991 as a point of forecast origin would have been problematic for several reasons. The forecast intervals failed to cover most of the real depression data: The percent coverage for males was 0.00 %, and that for females was a little more than 5.2 %. Secondly, 19 out of 31 observations were excluded from the estimation segment of the data, generating a convergence problem for estimation of our the initial full models.

The use of 1996 as a point of forecast origin for the percent coverage, MAE, and MAPE were better for both men and women than when we forecast from 1991 and 2005. When we attempted to forecast from 1996, 15 observations were excluded from the estimation segment, generating convergence problems for estimation of the initial full model for both males and females. For these reasons, we used 2005 as the point of forecast origin, by which we more easily circumvented the major intervening and potentially confounding events than if we use an earlier date.

Table 14: **Female depression forecast evaluations from 3 different points of origin**

Forecast origin year	Horizon length	Percent coverage	MSFE	RMSFE	MAE	MAPE
1991	19	5.263	0.013	0.115	0.099	69.375
1996	15	68.421	0.001	0.030	0.018	6.601
2005	5	60.000	0.006	0.076	0.061	16.312

11.6 Internal validity

Historically intervening variables are a source of challenge to the internal validity of a study [9, 5]. Therefore, we had to try to avoid the impact of the multiple political and economic crises of 2006-2010 by our temporal reversion to an earlier point of forecast origin, lest the universal impacts of these events induce systematic error from which even random selection may not completely protect [46, 209-210].

11.7 Statistical conclusion validity

With our sample size of 702, there were 339 males and 363 females. We obtained a large enough sample for us to have the statistical power of the analysis to detect small-to-medium effects for any of our regression analysis. Augmentation of regression effects does not appear to have impaired the filtering process. We also determined that our perceived health risk index had sufficient reliability before applying it in our analysis, shown in Table 2. For the state space models, we test the model assumptions and find our models to be well-specified by evidence of their fulfillment. With the high proportion of variance explained and general fulfillment of model assumptions, we believe our models to have reasonable statistical conclusion validity. In our models, the initial prelude period before 1986 serves as a partial self-control.

11.8 Generalizability

Because our sample was a random telephone sample of the residents in the two oblasts, it was a representative sample of the general Ukrainian public within the two oblasts surveyed. We randomly generated phone numbers that we attached to the area codes provided by the Ukrainian telephone company. From each household called, where we obtained permission for an interview, we interviewed only one respondent. Only after a separate auditing group confirmed that all of these responses were voluntarily provided without undue guidance by the interviewer, were the data uploaded to the Vovici corporation for initial dataset construction. Most previous studies were not completely representative samples, whereas the respondents selected for our sample were contacted only by randomized phone number generation. External validity is much more effectively assured by random selection of respondents, with sufficiently large samples, than by mere exclusion-inclusion criteria, propensity score matching, or post-stratification used in most case-control or other observational studies [46, 209-210],[37, 212-221].

We endeavored to circumvent major intervening variables that could confound the analysis of depression, whereas other studies did not. Bromet et al. (2011) have noted that it is all but impossible to disentangle these long-lasting effects[8]. Although we make no claim that we are able to overcome all possible potentially intervening variables, we employed an earlier point of forecast origin to circumvent incidents potentially associated with major surges of depression after 2005 (Figures 4 and 6). The evidence of the impact of these events were the surges in annual average reported depression levels, exceeding the 95% confidence limits of our 2005 to 2010 forecasts. The impact of historically intervening variables on an endogenous variable can undermine internal validity of a study. Temporal reversion to an earlier point of forecast origin avoids the impact on psycho-social depression by the gas disputes, political turmoil, and great economic recession of 2008 through 2010, so we could focus on the level and trends of depression following the accident at Chornobyl. Moreover, these multiple crises occurred one adjacent to the other with the great recession overlapping the 2009 gas dispute. One crisis could mask an overlapping crisis one or in the case of adjacent crises, the effects of one could smear the effects of the other. It would have been difficult to distinguish individual crises from this crisis patch. To circumnavigate this turmoil of potentially confounding variables and to convey a sense of what would have otherwise endured, we applied a form of a scenario forecast of what would have followed, had that intervening turmoil not taken place. This technique may be useful for averting such end-effect problems confronting forecasters.

11.9 Limitations

One limitation of this approach to circumvention is that other intervening events may have impacted our analysis. We do not make any claim to have eliminated all such effects. By employing an earlier point of forecast origin, we avoid having to model the political and economic intervening effects within the 2005 through 2010 forecast horizon. The potentially confounding events we avoid having to model are associated with the largest surge in depression evident in the time series displayed in Figures 4 and 6, since the Chornobyl accident. Without circumventing these intervening impacts, it would be extremely difficult if not impossible to analyze psycho-social effects of the nuclear accident at Chornobyl.

Another limitation is our dependence on the recollection of respondents to report substantial and significant changes in their levels of depression. The Chornobyl accident occurred in 1986 and, at the time of our interviews, we had no other way to obtain a full description of the psychological effects of that accident. Without such an analysis, there would be no empirical data on these effects. Nor would we have enough empirical evidence of how to deal with a nuclear incident, whether from an inadvertent accident or from a radiation dispersal device (RDD) or an improvised nuclear device (IND).

To combat problems of recall bias, we employed techniques to facilitate the memory of significant events. We employed simple units of measurement. We asked respondents to quantify responses on a percentage scale of 0 to 100. We dropped the lowest 5% so as to minimize confusion with moods swings and sadness. We employed periodization with simple salient temporal markers to facilitate recall, such as the year of Chornobyl, a period of of 1987 through 1996, and finally, the period since then until the end of 2009. We posed these questions in various ways to help respondents recall important events and we asked about associated medical diagnoses.

Because we began our point of forecast origin in 2005, we had only 26 years of data since the Chornobyl nuclear accident until we had to begin forecasting. Taking annual averages of variables meant that the number of time periods in our sample was not large. We needed to use a method of time series analysis that can be used with smaller nonstationary series. Therefore we restricted our covariates to a very small number of variables that we needed for testing variables related to our research questions. We could not pack our tests with a large number of socioeconomic variables for loss of degrees of freedom for testing. In a retrospective study age is constant and even when allowed to be time-varying, it did not prove to be a significant predictor in these models.

11.10 Directions for future research

The models we develop are merely a first step toward developing a more general depression model for persons in the area beyond the exclusionary zone but in the general vicinity of a nuclear event. Perhaps when other data are collected following other events, these results may be replicated there. It may also be possible to combine the data to perform a more general depression predictive model at some later point in time.

We could also consider the impulse indicator saturation, step indicator saturation techniques of Sir David Hendry, Søren Johansen, and Carlos Santos (2008) to identify structural breaks to model [31, 1-36] [15]. We could attempt super-saturation with broken trends or ultra-saturation suggested by Ericsson et. al. (2014) [18] with the inclusion of the foregoing plus interactions in future research.

Until then, this article may serve as a means of projecting what psycho-social depression may be expected of persons residing outside the exclusionary zone around the site of a nuclear incident.

References

[1] Anderson, B.D.O. and Moore, J.B. (1979) Optimal Filtering Mineola, NY: Dover Publications, 36.

[2] Apsimon, H. and Wilson, J. (1991) The Application of Numerical Models to Assess Dispersion and Deposition in the event of a Nuclear Accident *Journal of Forecasting(***10)**, 91-103.

[3] Balonov, M.I. (2007) The Chornobyl Forum: findings and recommendations *Journal of Environmental Radioactivity,* **96**, 6-12.

[4] BBC Ukraine historical timeline February 23, 2014 http://news.bbc.co.uk/2/hi/europe/country_profiles/1107869.stm

[5] Biernson, G. 1990 Optimal radar tracking systems New York: Wiley, 13, 413.

[6] Bromet, E. J. (2012) Mental Health Consequences of the Chernobyl disaster *Journal of Radiological Protection 32(1).*

[7] Bromet, E.J., Havenaar, J.M. (2007) Psychological and perceived health effects of the Chernobyl disaster *Health Physics 93(5).*

[8] E.J. Bromet, J.M. Havenaar y, L.T. Guey (2011) A 25 Year Retrospective Review of the Psychological Consequences of the Chernobyl Accident *Clinical Oncology, 23*, 297-305.

[9] Campbell, D. T. and Stanley, J. C. (1963) Experimental and Quasi-experimental design Chicago, Ill: Rand McNally Publishing, 7.

[10] The Chornobyl Forum: 2003-2005, D. Kinley III (editor). 2006 Chornobyl's Legacy: Health, Environmental, and Socio-economic Impacts and Recommendations to the Governments of Belarus, the Russian Federation, and Ukraine, 2nd revised edition Vienna, Au: International Atomic Energy Association, 10-20, 36-37.

[11] Commandeur, J.F. and Koopman, S. J. (2007) An Introduction to State Space Time Series Analysis Oxford, UK: Oxford University Press.

[12] DeCort, M., et al. (1998) Atlas of Caesium Deposition on Europe after the Chernobyl Accident Luxembourg: Office for Official Publications of the European Communities.

[13] deJong, P. (1988) The Likelihood for a State Space Model *Biometrika,* **(75)**, 165-169

[14] deJong, P. (1991) The Diffuse Kalman Filter *The Annals of Statistics, Vol. 19,***(2)**, 1073-1083.

[15] Doornik, J.A. and Hendry, D. F. 2013 Empirical Econometric Modeling PCGive 14, Volume 1 London, UK: Timberlake Consultants, Ltd., 215-234.

[16] Doornik, J.A. and Hendry, D. F. 2013 Econometric Analysis with Markov Regime Switching Models PCGive 14, Volume V London, UK: Timberlake Consultants, Ltd., 6-10.

[17] Durbin, J. and Koopman, S. J. (2001) Time Series Analysis by State Space Methods Oxford, UK: Oxford University Press, 115-120.

[18] Ericsson, N., Hood, S.B., Joutz, F. Sinclair, T., and Stekkler, H.O. Greenbook forecasts and the Business Cycle Presentation at 2014 Federal Forecasters Conference, Bureau of Labor Statistics, April 24, 2014.

[19] Ginzberg, H. (1993) *The Psychological Findings of the Chornobyl Accident: Findings from the International Atomic Energy Agency Study. Public Health Reports*, 108(**2**), 192-194.

[20] Godet, M. (1982) From forecasting to "La Prospective" - A new way of looking at futures *Journal of Forecasting (1)*, 293-201.

[21] Granger, C.W.J. and Newbold, P. (1974) Spurious regressions in Econometrics *Journal of Econometrics 2*, 111-120..

[22] Harvey, A. C. (1989) Forecasting, structural time series, and the Kalman Filter Cambridge, UK: Cambridge University Press, 268.

[23] Havenaar, J.M. and van den Brink, W. 1997 Psychological Factors Affecting Health After Toxicological Disasters *Clinical Psychology Review, 17(4)*, 365.

[24] Havenaar, J. M., Wilde, E. J., van den Bout, J., Drottz-Sj oberg, van den Brink, W. (2003) Perception of risk and subjective health among victims of the Chernobyl disaster *Social Science & Medicine, 56, Issue 3*, 569-572.

[25] Hyndman, R.J., Koehler, A.B., Ord, J.K., and Snyder, R.J. (2008) Forecasting with Exponential Smoothing: the state space approach New York, NY: Springer.

[26] Hyndman, R.J., Koehler, A.B., Snyder, R.J., and Grose, S. (2002) A state space framework for automatic forecasting using exponential smoothing methods *International Journal of forecasting, 18*, 439-454.

[27] Institute of Navigation (Accessed February 16, 2014) Navigation Museum *Apollo 8 Guidance and Navigation System.* https://www.ion.org/museum/item_view.cfm?cid=4scid=19iid=293

[28] Jacob, P., H.G. Paretzke, and H. Rosenbaum (1988) Organ doses from radionuclides on the ground. Part II. Non-trivial time dependences. *Health Physics, 1988. 55(1)*, 37-49.

[29] Jacob, P. et al. (1988) Organ doses from radionuclides on the ground. Part I. Simple time dependences. *Health Physics, 1988. 54(6)*, 617-33.

[30] Kahn, H. (1965) On Escalation: metaphors and scenarios New York: Praeger.

[31] Johansen, S. and Nielson, B. (2009) *An Analysis of Indicator Saturation Estimator as a robust regression estimator* in Castle, J.L. and Shepphard, N, eds. The Methodology and Practice of Econometrics, 1-26

[32] Koopman, S. J., Harvey, A.C., Doornik, J.A., and Shephard, N. Structural Time Series Analyser, Modeller, and Predictor STAMP 8.2 (2009) London, UK: Timberlake Consultants, Ltd , 175.

[33] Koopman, S. J., Shephard, N., and Doornik, J.A. (2008) SsfPack 3.0 (2008) London, UK: Timberlake Consultants, Ltd , 175.

[34] Leamer, Ed (1983) Let's take the con out of econometrics *American Economic Review, 73*, 31-43.

[35] Likhtarev, I. A. et. al. (1996) Effective doses due to external irradiation from the Chernobyl accident for different population groups of Ukraine. *Health Physics, 70(1)*, 87-97.

[36] Likhtarev, I. A. et al. (2002) Chernobyl accident: Retrospective and prospective estimates of external dose of the population of Ukraine. *Health Physics, 82(3)*, 290-303.

[37] Pedhazur, E. J. and Schmelkin, L. P. (1991) Measurement, Design, and Analysis Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 212-219.

[38] Protzman, F. (January 23, 1990) Upheaval in the East; East Germany Discloses Serious Accident at Nuclear Plant in 1976 http://www.nytimes.com/1990/01/23/world/upheaval-east-east-germany-discloses-serious-accident-nuclear-plant-1976.html Accessed 19 Feb. 2014.

[39] Perez Foster, R. and Goldstein, M.F. (2007) Chernobyl disaster sequela in recent immigrants to the United States from the former Soviet Union. *Journal of Immigrant and Minority Health, 9*, 115-124.

[40] Perez Foster, R. and Branovan, D.I. (2006) Surviving Chernobyl in America: Physical and mental health needs of recent FSU immigrants to the United States. *Journal of Immigrant and Refugee Studies, 4*, 97-118.

[41] Perez Foster, R. (2002) The long-term mental health effects of nuclear trauma in recent Russian immigrants to the United States. *American Journal of Orthopsychiatry, 72*, 492-504.

[42] Nunnelly, J.C. (1978) Psychometric Theory New York, NY: McGraw Hill, 226.

[43] Rodgers, R. M. (2003) Applied Mathematics in Integrated Navigation Systems 2nd. ed. Reston, VA: American Institute of Aeronautics and Astronautics, xv, 206, 213.

[44] Saenko, V., Ivanov, V, Tsyb, A., Bogdanova, T., Tronko, M., Yu, D. and Yamashita, S. (2011) The Chernobyl Accident and its Consequences Clinical Oncology, 23, 234-243.

[45] Saito, K. et al. (1996) Calculation of organ doses from environmental gamma rays using human phantoms and Monte Carlo methods. Part 1: Monoenergetic sources and natural radionuclides in the ground Gesellschaft fur Strahlen- und Umweltforschung.

[46] Schuemie, M.J., Ryan, P.B., DuMouchel, W., Sucharrd, M. A., and Madigan, D. (2014) Interpreting observational studies: why empirical calibration is needed to correct p-values. *Statistics in Medicine, 33*, 209-218.

[47] Sperb Leite, M. and Roper, L. David (1988) *The Goiânia Radiation Incident A Failure of Science and Society* http://arts.bev.net/roperldavid/gri.htmConclusion Accessed 19 Feb. 2014.

[48] Tashman, L. (2000) Out-of-sample forecast evaluation *International Journal of Forecasting, 16*, 437-450.

[49] Time and Life Pictures of the Worst Nuclear accidents (Accessed 19 Feb 2014) Time Photos http://content.time.com/time/photogallery0,29307,1887705_2255451,00.html

[50] Wack, P. (1985) Scenarios: Shooting the rapids Harvard Business Review, September-October, 73-89.

[51] WISE (2010) Criticality accident at Tokai nuclear fuel plant(Japan) World Information Source on Energy Uranium Project http://www.wise-uranium.org/eftokc.htmlWORLD Accessed 19 Feb 2014.

[52] Yaffee, R. A., Nikolopoulos, K., Reilly, D.P., Crone, S.F., Wagoner, K.D., Douglass, R.J., Amman, B.R., Ksiazek, T.G., and Mills, J.M. (2008) An Experiment in Epidemiological Forecasting: A Comparison of forecast accuracies of different methods of forecasting Deer Mouse Population Density in Montana A report on forecasting research done with the Centers for Disease Control and Prevention, Atlanta, Ga. Presentation at George Washington University, Center for Economic Research, December 1, 2011. http://research.columbian.gwu.edu/cer/research/forecasting/pastpresentations, 16-19.