Quantifying Change Over Time: Interpreting Time-varying Effects In Duration Analyses

Constantin Ruhe[1,2]

[1] Researcher, German Development Institute/Deutsches Institut für Entwicklungspolitik (DIE), 53113 Bonn, Germany
[2] Associated Fellow, Zukunftskolleg/Department of Political and Administrative Science, University of Konstanz, 78457 Konstanz, Germany. Email: Constantin.Ruhe@die-gdi.de

Abstract

Duration analyses in political science often model nonproportional hazards through interactions with analysis time. To facilitate their interpretation, methodologists have proposed methods to visualize time-varying coefficients or hazard ratios. While these techniques are a useful, initial postestimation step, I argue that they are insufficient to identify the overall impact of a time-varying effect and may lead to faulty inference when a coefficient changes its sign. I show how even significant changes of a coefficient's sign do not imply that the overall effect is reversed over time. In order to enable a correct interpretation of time-varying effects in this context, researchers should visualize their results with survivor functions. I outline how survivor functions are calculated for models with time-varying effects and demonstrate the need for such a nuanced interpretation using the prominent finding of a time-varying effect of mediation on interstate conflict. The reanalysis of the data using the proposed visualization methods indicates that the conclusions of earlier mediation research are misleading. The example highlights how survivor functions are an essential tool to clarify the ambiguity inherent in time-varying coefficients in event history models.

Keywords: duration models, Cox model, proportional hazards, time-varying effects, survivor function, mediation

1 Introduction

Duration analyses in political science frequently examine data over long periods of time. Yet, as time passes, the effects of variables often change. In the widely used Cox Proportional Hazards model, this phenomenon will cause the well-known violation of the proportional hazards assumption (Cox 1972; Box-Steffensmeier and Zorn 2001). Directly modeling the time-varying effect through interactions with some function of analysis time can solve this problem. It also enables to investigate the effect, if the time-varying effect is of substantial theoretical interest (Box-Steffensmeier, Reiter, and Zorn 2003). While this modeling approach is easy to implement, the substantive interpretation is not straightforward. Hence, political scientists have developed techniques to ensure that time-varying hazard ratios are visualized correctly (Licht 2011; Gandrud 2015). However, I demonstrate in this paper that these existing techniques only describe a variable's instantaneous and multiplicative effect. Time-varying hazard ratios provide no indication about the absolute change in risk and can be very ambiguous when the overall, cumulative effect is of interest. In these circumstances, a clear interpretation requires additional calculations. This is especially true if the time-varying effect implies that the coefficient significantly reverses its sign. I show how researchers can eliminate this ambiguity and graphically

analyze their results using survivor functions to support valid conclusions on how strongly an effect changes over time. To demonstrate how appropriate visualizations using survivor functions may clarify and even change substantive conclusions, I reevaluate the time-varying effect of third-party mediation in interstate conflict (Beardsley 2008, 2011).

Throughout the paper, I mainly focus on the Cox Proportional Hazards model, which is often the first choice for applied duration modeling in political science (Cox 1972; Box-Steffensmeier and Jones 2004). Nevertheless, the general implications for time-varying effects are also valid for parametric models which assume proportional hazards. The Cox model's popularity in political science stems from the fact that it does not require an *a priori* assumption about the distribution of the baseline hazard. However, the unknown baseline may make a substantive interpretation of time-varying effects very complex. While hazard ratios provide an intuitive interpretation for a basic model with constant effects, time-varying effects can be highly misleading (cf. Royston and Parmar 2011). Although political science has proposed good solutions to visualize how the hazard ratio varies with time (Licht 2011; Gandrud 2015), I demonstrate that time-varying hazard ratios or relative hazards are quite ambiguous and leave room for very different substantive interpretations about a variable's overall effect. In fact, a significant change in a coefficient's sign can imply three different substantive conclusions: First, a variable could decrease/increase the duration or the probability of an event, but after some time, the variable begins to have the opposite effect. Second, a variable might decrease/increase the duration or the probability of an event, but this effect disappears at some point. Third, a variable might permanently decrease/increase the duration or the probability of an event, but the effect simply becomes somewhat smaller over time.

The central problem why time-varying hazard ratios or relative hazards are not sufficient to tell these effects apart arises from two issues: First, hazard ratios quantify merely a multiplicative change relative to some hazard rate. Second, even if the hazard rate is known, it is difficult to interpret because it is a conditional quantity which describes an instantaneous rate of failure, given that an event has not yet occurred (Box-Steffensmeier and Jones 2004, 14). If the proportional hazard assumption applies and a covariate affects the hazard by a constant factor, i.e., by the hazard ratio, the conditionality and the instantaneous nature are not that relevant, since the covariate simply changes the overall level of the hazard rate by the same factor at any given point in time. In contrast, with time-varying effects the substantive meaning of a time-varying hazard ratio depends both on the time-varying hazard ratio itself as well as the effect of other, potentially time-varying covariates and the baseline hazard (cf. Putter *et al.* 2005). This is the case because a time-varying effect with a change in sign implies that a variable causes first an increased or decreased instantaneous probability of failure, while later on, the opposite effect occurs. Depending on how much risk is accumulated or avoided at early stages of the study period compared to the opposite effect at later stages, the total effect of a variable can change, disappear or become merely somewhat smaller.

In this paper, I show how survival functions are able to provide the information to tell these effects apart and provide a very intuitive method to interpret the overall influence of a time-varying effect. Since survival functions provide the model's unconditional predicted probability of survival over time for specific covariate values, they are an easy and unambiguous method to communicate the overall impact of a time-varying effect even to audiences with limited statistical training (Putter *et al.* 2005). In this way, they can safeguard against inferential mistakes among the broader readership of social science research. While it is tedious to calculate survival functions for models with time-varying coefficients manually, applied researchers no longer face this obstacle, since these calculations have now been automated both in R and SAS (Thomas and Reyes 2014) as well as Stata (Ruhe 2016).

In the following sections, I discuss the complex interpretation of time-varying effects in duration analyses. Based on this discussion, I describe how researchers can use survival functions to

effectively visualize the implication of time-varying effects. I apply this approach to an example of immense policy relevance, the time-varying effect of third-party mediation (cf. Beardsley 2008, 2011). In the application, I demonstrate how an appropriate visualization of time-varying effects can substantively clarify and even change the policy implication. The replication highlights that, contrary to earlier interpretations, the time-varying effect of mediation does not suggest a problematic long-term effect on postconflict stability. Quite the contrary, mediation appears to correlate with a substantively higher chance of several years of peace. Despite a time-varying effect, which significantly reverses its sign, there is no indication that mediation creates adverse long-term effects. Beyond the substantive relevance for international relations research, the application demonstrates how survivor functions enable researchers to visualize and interpret time-varying effects in duration models intuitively, regardless of their substantive research interest.

2 Nonproportional Hazards in Political Science and Their Interpretation

2.1 The need to clarify the relevant quantity of interest

Time-varying effects are found in all subfields of political science (cf. Licht 2011; Box-Steffensmeier, Reiter, and Zorn 2003; Chiozza and Goemans 2004; Allen 2005; Golub 2007; Murillo and Martínez-Gallardo 2007; Beardsley 2008, 2011; Zhelyazkova and Torenvlied 2009; Hale 2015; Grewal and Voeten 2015). While existing methods to interpret time-varying effects enable to describe a variable's instantaneous effect (cf. Golub and Steunenberg 2007; Licht 2011), they do not allow clear statements about the change in effect magnitude and the overall effect of a variable over time. I show that this is unfortunate, since a time-varying effect can significantly change its sign, but still produce a positive or negative overall effect. Due to this fact, researchers need to clarify, whether their research question requires a focus on the instantaneous or the overall, i.e., the cumulative effect of a variable. In order to provide social scientists with a tool to describe the cumulative effect of such variables, I introduce survival functions for time-varying effects.

Social science research of duration processes can have very different aims and the relevant quantity of interest depends on the research question. For example, a theory might predict how variables affect the duration T of some process. Alternatively, it could formulate hypotheses about changes in the probability that the process continues or that it ends. In survival analysis, the probability that a process continues until some time t is described by the survival function $S(t) = Pr(T > t)$.[1] Since most models estimate how a variable affects the hazard rate, a theory can also describe how a variable affects the immediate risk at a specific point in time (cf. Box-Steffensmeier and Jones 2004).

A very general theory might simply predict that a variable increases or decreases the duration, the probability of an event or the hazard rate. If the effect is constant over time, the quantity of interest used to test the hypothesis does not matter much. If a variable with a constant effect increases the hazard rate, the hazard rate will be higher at any time. This also corresponds to an overall higher probability of failure, a lower probability of survival as well as a shorter average duration. However, if a researcher subsequently detects nonproportional hazards, which imply a time-varying effect, the quantity of interest matters. In this context, it becomes essential to determine whether the researcher is interested in the overall effect which the variable creates over time, i.e., the cumulative effect, or whether the interest lies with the instantaneous effect.

Cumulative effects will be of particular interest for variables, which remain constant over a longer period or even the entire duration. Variables such as whether a conflict ended in a stalemate do not change once the conflict ended, although their influence on the outcome might evolve with time (cf. Box-Steffensmeier, Reiter, and Zorn 2003). Similarly, regime type will often be

---

1 The survival function also allows to predict median duration (survival time) or the average survival time up to a certain time (cf. Cleves *et al.* 2010; Royston and Parmar 2011).

constant throughout the duration or, given that there is a change at some point, it will persist for a longer time after the change occurred (cf. Chiozza and Goemans 2004). For more quickly or frequently changing time-varying covariates, both instantaneous and cumulative effects may be of interest to researchers. However, regardless of their research interest, researchers should always consider the assumptions associated with time-varying *covariates* in duration analyses, regardless of whether these variables display time-varying *effects* (see Box-Steffensmeier and Jones 2004, 95ff.).

The difference between cumulative and instantaneous effects becomes clear with the example of education. Let us assume that a researcher postulates that job training increases income. During the research process, it becomes clear that going to school decreases the immediate earnings while increasing future wages. In this context, the researcher might now study how the additional training affects a person's earnings at different times in their life. Alternatively, the researcher could analyze if the training increases lifetime earnings. Let us assume that the hypothetical job training reduces earnings to almost zero for 3 years, while increasing wages by 10 percent after about 5 years. While this information answers how the training affects wages at specific points in time, it does not provide enough information to answer the question whether the training pays off over an entire career. This question depends on the monthly wage that was lost and how high the total amount of a 10 percent increase in wages actually is. It further depends on how long participants will continue to work after completing the training. Based on the percentage changes over time alone, it is impossible to say whether the training pays off, whether the losses and gains even out or whether a low salary level and a short remaining time to work are unable to make up for the income lost during the training period.

Similar to the education example, a variable in a duration model may, e.g., decrease the immediate risk of failure early on, but increase the instantaneous risk at a later time. As in the lifetime income example, the theoretical prediction could be that the variable is associated with an overall lower probability of an event. I show below that the hypothesis could still be true, despite a time-varying effect, which changes its sign. Hence, to interpret the substantive implication of time-varying effects, researchers need to clarify whether they are interested in the instantaneous or the cumulative effect. Since existing methodologies to interpret time-varying effects describe only the instantaneous, multiplicative effect of the variable, I introduce survival functions for time-varying coefficients which enable to visualize the cumulative effect as well as its absolute magnitude.

2.2 Expanding the interpretation of time-varying effects

The most commonly used duration models in political science assume proportional hazards which imply that variables have a constant effect over time (Box-Steffensmeier and Jones 2004). Since a violation of this assumption can undermine the validity of the model, political scientists have developed helpful strategies to detect and adequately model nonproportional hazards (Box-Steffensmeier and Zorn 2001; Keele 2010; Park and Hendry 2015). Since even adequately modeled nonproportional hazards are not as easy to interpret as proportional hazards, a second strand of research has developed tools to calculate meaningful quantities of interest (Golub and Steunenberg 2007; Licht 2011; Gandrud 2015). In this paper, I add to the latter part of the literature and discuss how existing interpretation techniques, such as time-varying hazard ratios or relative hazards can be very ambiguous and, in the worst case, may result in misleading inference about the substantive effects. I introduce survival functions for time-varying effects as a suitable technique in how researchers can reduce this ambiguity and visualize the implications of their results more clearly.

Before nonproportional hazards can be interpreted, however, they need to be identified and adequately modeled. Thereby, it is important to keep in mind that not all violations of

---
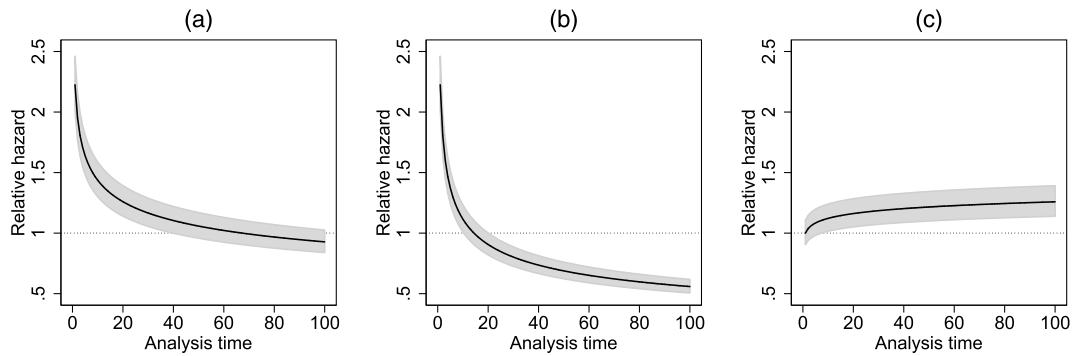
Figure 1. Potential patterns of time-varying hazard ratios or relative hazards.

the proportional hazards assumption indicate time-varying effects; these can also arise from an incorrectly specified functional form (Keele 2010). If nonproportional hazards are present even with a correct functional form, interactions with time are an easy approach to model a time-varying effect on the hazard of observing an event (Box-Steffensmeier and Zorn 2001). In the widely used Cox model, this leads to the following model: Let $h_0(t)$ be an unspecified baseline hazard function of observing the event of interest, which can take on any form. If we model a time-varying effect through an interaction with time, the hazard function for an observation i is then asserted to be

$$h(t|x_i) = h_0(t)e^{x_{1i}\beta_1 + x_{2i}(\beta_2 + \beta_3 f(t))}, \tag{1}$$

whereby the effect of x_1 is assumed to be constant while the effect of x_2 is allowed to vary with some function of analysis time (cf. Box-Steffensmeier and Zorn 2001).[2] If the model is a discrete duration model, e.g., a Logit or Probit model with time dependence (cf. Beck, Katz, and Tucker 1998; Box-Steffensmeier and Jones 2004), nonproportional hazards can be modeled through a similar interaction with time (Carter and Signorino 2010a).

Hence, time-varying effects are easily introduced in a model. Unfortunately, however, the substantive meaning of a time-varying effect is not straightforward. First, the interaction effect needs to be interpreted correctly (cf. Brambor, Clark, and Golder 2005). Golub and Steunenberg (2007) as well as Licht (2011) show for the widely used Cox model how the combined coefficient can be used to calculate time-varying hazard ratios as well as relative hazards. If visualized correctly, these techniques indicate how a variable's effect on the hazard rate changes with time. It also highlights when these effects are significant.

If we assume the commonly estimated logarithmic effect $(\beta_2 + \beta_3 \times \ln(t))$, several patterns can occur. Figure 1 displays how these patterns might look like if the corresponding hazard ratios or relative hazards are visualized using the method proposed by Licht (2011): First, the effect may decrease in size (and possibly become insignificant at some point), as depicted in (a). Second, the effect might decrease in size and eventually significantly reverse its sign (see (b)). Finally, as shown in (c), the effect size could actually increase and possibly become only significant after a certain time.[3]

Although this type of visualization is sufficient to highlight the pattern with which the instantaneous effect changes, I show in this paper that it does not provide a clear indication

---

2 There are different choices for the functional form of the effect's change over time. In many political science applications the natural logarithm of time $\ln(t)$ is used (cf. Box-Steffensmeier and Zorn 2001), but many others are plausible and potentially more appropriate. Therefore, the functional form needs to be chosen carefully (Park and Hendry 2015).

3 Obviously, all scenarios could also be depicted for (initially) negative effects, that is with reversed sign. With nonmonotonic transformations, further effect patterns are possible.

about the overall effect over time as well as its changing magnitude. This is due to the fact that the instantaneous effect of a variable might be outweighed by the different earlier effect which a variable created. For example, the risk, which was avoided early on, might outweigh the increased risk at later points in time.

I discuss how visualizing the results with survivor functions can give a good intuition of the effect magnitude in the data and generate predictions for substantively interesting scenarios. In panels (a) and (c) survival functions provide an intuitive interpretation of the overall effect, in addition to relative hazards. For example, a survival function can show, whether an effect of type (a) still causes a higher/lower probability of survival, even after the variable has lost its immediate influence. In scenario (b), however, i.e., when the estimated effect reverses its sign, survival functions are a crucial step for an unambiguous interpretation. The necessity for a survival function in scenario (b) arises from the fact that a significant change in a coefficient's sign can support three different substantive conclusions about the overall effect: First, the variable might decrease/increase the duration or the probability of an event, but after some time the variable begins to have the opposite effect. Second, the variable could decrease/increase the duration or the probability of an event, but this effect disappears at some point; third, the variable might permanently decrease/increase the duration or the probability of an event, but the magnitude of the effect becomes somewhat smaller over time.

Hence, scenario (b) entails a lot of ambiguity. It implies that a simple hypothesis like "higher values of X increase the duration of Y" can still be valid, even if the estimated time-varying effect significantly reverses its sign. This ambiguity ensues because both relative hazards and hazard ratios describe a multiplicative change of an unspecified baseline as well as the fact that this baseline is an instantaneous rate of failure, given that the event has not yet occurred. I discuss the importance of both of these factors in detail in the next two sections and describe how survival functions incorporate them. Since the approaches outlined by Golub and Steunenberg (2007) as well as Licht (2011) focus on the multiplicative, instantaneous effect, they only allow to describe a time-varying effect as e.g., type (b), but they do not allow to describe the substantive overall impact of such a time-varying effect. The visualization using survival functions proposed in this paper overcomes these limitations.

Discrete duration models (also used as binary time-series-cross-section models) are another form of duration model frequently used in political science (cf. Beck, Katz, and Tucker 1998; Box-Steffensmeier and Jones 2004). Carter and Signorino (2010a) show that nonproportional hazards can be easily modeled and visualized in these models, since the baseline hazard is estimated using a flexible function of time. Hence, the magnitude at a given point in time can be calculated. In fact, if events occur repeatedly, Williams (2016) shows that changes in the probability of an event can also affect the long-term effect of additional future events at a certain point in time. If a variables effect changes over time, these nonproportional hazards also need to be modeled to accurately estimate possible long-term effects (Williams 2016). Hence, discrete duration models are therefore more easily able to estimate the magnitude of the change in the hazard rate, or more precisely the hazard probability at a given point in time. However, this quantity of interest remains uninformative about the total effect over time, since it is also an instantaneous failure rate, given that the event has not yet occurred. This leaves the same ambiguity regarding the overall implication of time-varying effects of type (b). Again, survival functions are a suitable tool to dissipate this ambiguity.

I use the example of a time-dependent effect of mediation in international crises to highlight the two central aspects which cause this ambiguity (cf. Putter *et al.* 2005):[4]

---

4 These are particularly relevant for the Cox Proportional Hazards model as well as parametric models, which make the proportional hazard assumption, e.g., the Weibull model or discrete duration models.

1. The context determines the magnitude of a time-varying effect. This context consists of the effect of other covariates, regardless of whether they are constant or time-varying, as well as the baseline hazard. Without knowledge of the (potentially time-varying) baseline, which an effect changes multiplicatively, it is not possible to describe the substantive implication of a time-varying effect.

2. Even if the values of other covariates and the baseline hazard are taken into account, an analysis of a hazard rate requires care. The hazard rate describes the instantaneous risk of failure at a point in time, given that the event has not yet occurred. This implies that the substantive importance of long-term effects depends on earlier short-term effects.

Below, I discuss these points and highlight how survival functions can help to overcome these problems and enable researchers to intuitively visualize time-varying effects.

3   The Importance of the Baseline

As described above, the substantive interpretation of a time-varying effect can heavily depend on the context. Time-varying hazard ratios describe how, at a specific time, some baseline value is increased or decreased multiplicatively by a variable. This means that the magnitude of the change is determined by this baseline. In turn, the baseline depends on the values and effects of other variables in the model as well as the baseline hazard rate. Thereby, the baseline hazard rate captures every remaining process that the model does not explain systematically based on independent variables.

Thus, the baseline hazard describes how the average risk, which is not explained systematically, evolves over time. Since the baseline hazard may therefore simply be "a statement about omitted variables" and consequently change with the model, the question whether the baseline hazard should be interpreted has caused some controversy (cf. Beck 2010, 294). Nevertheless, others have argued that until a better model can be constructed, the baseline hazard is a substantive part of the model, which contains important information about the underlying data (cf. Carter and Signorino 2010b, 296f.). Although I generally agree with the perspective by Beck (2010), the absolute magnitude as well as the cumulative effect of a variable in a given dataset are only identifiable if we use the information about the underlying data provided by the baseline hazard. A second aspect reinforces this perspective: Even when a duration process is perfectly understood and modeled, leaving only a flat baseline hazard, time-varying covariates or other covariates with time-varying effects can lead to changes in risk over time. Hence, in this case, the baseline, not the baseline hazard rate, increases or decreases over time and this changes the overall magnitude of a time-varying effect and potentially even its substantive meaning. Below, I therefore use the word baseline to highlight that there are multiple possible causes for changes in this baseline.

Most parametric models assume a specific functional form for the baseline hazard. However, the Cox Proportional Hazard model allows to estimate the effect of a variable on the hazard rate of observing an event at time t without any specification of the functional form of the baseline hazard rate. This flexibility of the semiparametric Cox model has led to the popularity of the Cox model in political science (Box-Steffensmeier and Jones 2004).[5] At the same time, however, not knowing the baseline makes any substantive interpretation of a time-varying effect challenging, especially if the coefficient reverses its sign.

To understand how the baseline is important to assess the substantive meaning of a time-varying effect of type (b) in Figure 1 and highlight the limitation of hazard ratios or relative hazards in this context, it is important to review the interpretation of the coefficient in a Cox model. Due to the model's nonlinearity and since the baseline hazard rate $h_0(t)$ is left unspecified, the coefficients

---

5   Similarly, discrete duration models estimate the baseline hazard with flexible functions of time, such as splines or polynomials (cf. Beck, Katz, and Tucker 1998; Carter and Signorino 2010a).

themselves have little meaning. Given this limited information, hazard ratios, the exponentiated coefficients, are the most intuitive interpretation of the estimated coefficients. They express the multiplicative change in the hazard rate for a one-unit change in the predictor variable, ceteris paribus (Box-Steffensmeier and Jones 2004). Hence, a one-unit change in x_1 in Equation (1) would change the unobserved hazard rate by the factor e^{β_1} at any given time. However, in contrast to a constant effect, a time-varying hazard ratio is not nearly as intuitive. Aside from the care which should be devoted to an interpretation of interaction effects (cf. Brambor, Clark, and Golder 2005), the difficult interpretation of this relative risk measure arises from the unknown value and shape of the baseline which determines how large the absolute change in risk is at various points in time. Without knowledge of the baseline, the risk of observing the event at a certain point in time remains unknown (cf. Putter *et al.* 2005). The effect of the predictor variables can therefore only be interpreted as shifts in the unknown hazard rate.

Figure 2 visualizes this difficulty. For simplicity, we can think of this example as modeling the risk of acquiring a disease. Assume a first scenario with a baseline which is initially very high, but quickly falls to a very low level.[6] If in this context a treatment x would initially lead to a substantive decrease in the very high hazard, this would imply a drastic decrease in risk. Assume further that, due to a time-varying effect, x more than doubles the hazard rate after several years (see panel (b)). How substantive these short- and long-term effects are essentially depends on the baseline which is altered by variable x. In scenario 1, a late increase in relative risk would be reasonably small, since the overall hazard rate at that point is very low (see panel (c)). Consequently, with this hypothetical hazard rate, the treatment x might still be a good option, despite the time-varying effect. On the other hand, consider the second scenario with a very different, strongly monotonically increasing baseline. Panel (c) shows that with a constantly increasing baseline, a hypothetical time-varying effect of treatment x implies a substantially elevated hazard rate at later points in time. In the supplementary information, I provide a further example which highlights that even proportional changes in a flat hazard rate can affect the cumulative effect of a time-varying hazard ratio.

The example highlights how context-dependent the actual magnitude of a time-varying effect can be. In scenario 1, we would probably conclude that the overall treatment effect of x is beneficial because it decreases the risk at a time when the risk is very high, while the increased hazard rate of treated people that remain healthy is neglectable. In contrast, scenario 2 is substantively more ambiguous. If the model contains more than one time-varying effect, this problem becomes even more pronounced. Depending on the value of these variables, the hazard rate may be increasing or decreasing. Consequently, the same time-varying hazard ratio might imply different substantive effects, given alternative values of the other variables with time-varying effects (cf. Putter *et al.* 2005).

Hence, hazard ratios are not a sufficient way to describe the substantive meaning of a time-varying effect on failure risk. Without knowledge of the baseline, the magnitude of a time-varying effect remains unclear. Nevertheless, this is not to say that researchers should not graph time-varying hazards ratios. Such a graphical analysis is very important to describe whether and how an effect changes with time and whether any changes are statistically significant (Licht 2011). However, these plots are not a good indicator of the magnitude of an effect. For this, we need to know the baseline. Even further steps are needed to enable a good interpretation of the overall effect of a variable.

---

6   Again, as described above, the decline in the hazard rate does not have to be due to a declining baseline hazard rate. This scenario could also arise from a flat baseline hazard (or any other shape) and a time-varying effect of an additional independent variable, which multiplicatively changes the baseline hazard and leads to the overall pattern of a declining hazard rate. It can also be due to time-varying covariates with proportional hazards.

(a) Baseline

(b) Hazard ratio of variable x

(c) Hazard rates for different values of x
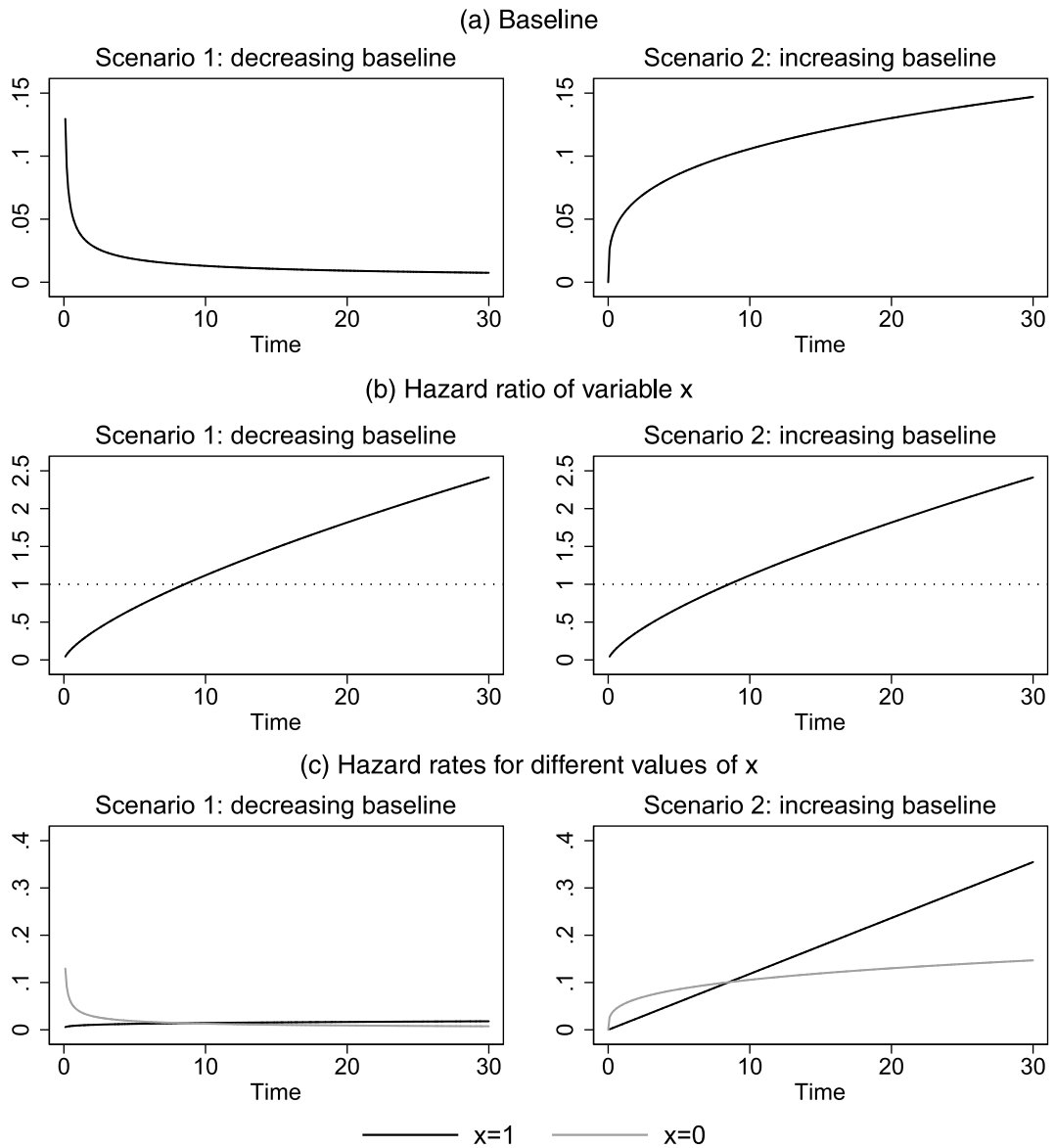
Figure 2. Same hazard ratio, different conclusion: The magnitude and substantive importance of an effect varies with the overall shape of the baseline. Results for two scenarios with the same time-varying effect of variable x, but different Weibull distributed baseline hazards.

4 Beyond the Hazard Rate

While plotting the hazard rate gives an intuition into how relevant changes in an effect may be, this option is not available in the Cox model, since it provides no direct estimate of this function. However, even with parametric models, Scenario 2 shows that plotting the scenario-specific hazard is also quite unintuitive and gives little insights about the cumulative effect. This becomes especially apparent if we consider the substantive meaning of the hazard rate. One can think of the hazard as the instantaneous rate of failure, conditional on the fact that an event has not occurred up to this point in time. It can be described formally as follows (cf. Box-Steffensmeier and Jones 2004, 14):

$$h(t|x_i) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t \mid T \geq t, x_i)}{\Delta t}. \tag{2}$$

The instantaneous nature and the conditionality, however, are a crucial complication when interpreting the cumulative impact of a time-varying effect. To highlight why hazard rates are often not sufficient, we can assume that Figure 2 reports the finding of a randomized clinical trial. The purpose of the study is to evaluate the effect of the treatment. For the instantaneous effect, we are simply interested in the risk of acquiring the illness at a specific point in time, given that a patient has remained healthy and given the treatment choice. In this case, the hazard rate would be sufficient. Nevertheless, it is important to remember what is being compared at a late point in the study period. Despite randomization, the groups might no longer be identical toward the end of the study. Assume that there are an equal proportion of patients with good and with poor health in both the treatment and in the control group. At the start, the groups are identical, except for the treatment. Assume further that the treatment reduces the risk of sickness initially, but the effect disappears quickly. In the treatment group, the patients will be stabilized as long as the treatment has an effect. Once the treatment loses its effect, these weak patients will most likely start to become sick. In the control group, the weak patients catch the disease very quickly because they are not protected by the treatment. Hence, they are no longer in the sample. If we now compare the treatment and control group at this late point in time, we compare a treatment group in which many weak cases remain against a control group, which consists mostly of patients with good health. If the treatment now loses its effect and the weak cases start to get sick, we will naturally see a higher rate of infection in the treated group than in the control group. This is because we are comparing a treated group, which still contains strong and weak patients against a control group, which, at this point in time, consists only of strong patients.

If our main research interest is to evaluate the cumulative effect, calculating the magnitude of a time-varying hazard ratio by multiplying it with the baseline is not sufficient.[7] It is only an intermediate step, since it provides the risk of illness at a given point in time, given that a patient is still healthy. For the cumulative effect, we would like to know the probability with which a patient remains healthy up to a certain point, depending on the treatment choice. A time-varying hazard rate in itself does not provide clear evidence whether this hypothesis is true or false. In fact, it is possible that a higher proportion of patients with treatment x will remain healthy, even long after the hazard rates crossed. In this case, a simple hypothesis that x increases or decreases the duration to an event is still valid, despite a time-varying hazard ratio. On the other hand, the time-varying hazard ratio can also imply that a variable's effect is reversed after some time. Whether this is the case depends on how much instantaneous risk was avoided compared to the control group and how strongly the relative risk changes later on.

Fortunately, survival analysis provides a tool to examine how many units have not yet experienced an event at a given point in time: Survivor functions. These functions estimate the proportion of cases which have not (yet) failed at a certain point in time (Box-Steffensmeier and Jones 2004).[8] Duration models, such as the Cox model can be used to estimate these quantities of interest, but the calculations are not straightforward in the presence of time-varying effects (cf. Putter *et al.* 2005). I discuss this in detail below. However, in the motivating example described in Figure 2, the respective survivor functions are easily calculated analytically. Figure 3 plots the results for both scenarios from Figure 2. In the first scenario, the time-varying effect leads to a substantively higher probability of survival, which eventually converges to the same level toward the end of the analysis time. However, in the second scenario, the time-varying effect leads to a higher probability of survival during the first half of the analysis time, and a lower probability of survival thereafter. If a physician faces the first scenario, the results clearly indicate that the treatment has a beneficial effect for some time, before this effect eventually disappears. The

---

7   Moreover, this step is not available in the Cox model.
8   Cumulative hazard functions are an alternative statistic to describe time-varying effects, since they plot the total risk of failure which has been accumulated up to a certain point.

(d) Survival estimates for different values of x

Scenario 1: decreasing baseline

Scenario 2: increasing baseline

x=1   x=0
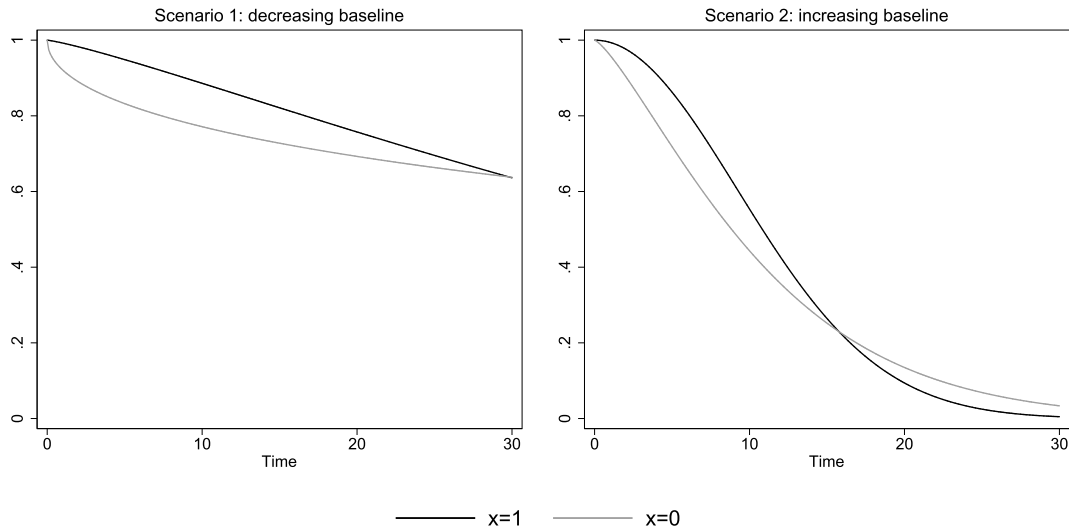
Figure 3. Survival estimates for scenarios presented in Figure 2.

second scenario is substantively more ambiguous, since the physician actually faces a trade-off between increasing short-term survival at the risk of long-term survival with treatment allocation.

The example highlights how different the cumulative implication of a time-varying effect can be. Depending on the baseline, crossing hazard rates could imply an impact, which is reversed over time, convergence between groups or even merely a slight decrease of a persistent difference between groups. Hence, to describe the overall effect, crossing hazard rates are a similarly ambiguous outcome as time-varying hazard ratios. Survivor or cumulative hazard functions are better suited to identify whether a time-varying effect reverses its impact over time.

5 Survival Functions for Time-varying Effects

How can a covariate-specific survival function be estimated in the commonly used Cox model? If the proportional hazards assumption holds, the survivor function for different covariate values can easily be calculated. Based on a hazard rate similar to (1), but without a time-varying effect, we can calculate the cumulative hazard function (Kalbfleisch and Prentice 2002; Cleves *et al.* 2010):

$$
\begin{aligned}
H(t|x_i) &= \int_0^t h(u|x_i)\, du \\
&= \int_0^t e^{x_i\beta} h_0(u)\, du \\
&= e^{x_i\beta} \int_0^t h_0(u)\, du \\
&= e^{x_i\beta} H_0(t).
\end{aligned}
\tag{3}
$$

Based on the cumulative hazard function with proportional hazards, we get the following survival function:

$$
\begin{aligned}
S(t|x_i) &= e^{-H(t|x_i)} \\
&= e^{-e^{x_i\beta} H_0(t)} \\
&= (e^{-H_0(t)})^{e^{x_i\beta}} \\
&= S_0(t)^{e^{x_i\beta}}
\end{aligned}
\tag{4}
$$

(Kalbfleisch and Prentice 2002; Cleves *et al.* 2010). Consequently, given proportional hazards, i.e., in the absence of time-varying effects, the survivor function for different scenarios can easily be calculated using the baseline survivor function $S_0(t)$ as well as the estimated coefficients. Since all statistical packages allow predicting the baseline survivor function from estimated Cox models, these calculations are easily implemented. Furthermore, all statistical packages provide automated tools to calculate covariate-specific survivor functions.

If we no longer assume proportional hazards and model a time-varying effect through an interaction with time, the calculation is not as simple. Due to the interaction with time, the linear combination of predictors $x\beta$ now changes to $x\beta(t)$, which is a function of time and remains in the integral:

$$H(t|x_i) = \int_0^t e^{x_i\beta(u)} h_0(u)\, du \tag{5}$$

(cf. Thomas and Reyes 2014). Moreover, since h_0 is not directly estimated in the Cox model, the calculation of the cumulative hazard rate and the survivor function are not directly available. Nevertheless, the model does provide estimates of the baseline cumulative hazard function as well as the baseline survivor function. These estimates are based on the information gained at each failure time and thus do not give an estimate of a smooth function, which leads to the familiar, jagged step functions. Each failure time thereby provides an estimate of the risk at that point in time: the hazard component. Based on the hazard component and the estimated coefficients, both the cumulative hazard as well as the survival functions can be calculated (see Kalbfleisch and Prentice 2002, 114ff.).

Based on the grouped relative risk model described in Kalbfleisch and Prentice (2002, 47f.), the hazard at failure time t_j and given covariates with time-varying effects $x\beta(t)$[9] can be estimated as

$$\Delta H(t_j|x_{i(t)}) = 1 - (1 - \Delta H_0(t_j))^{e^{x_{i(t_j)}\beta(t_j)}}, \tag{6}$$

whereby $\Delta H_0(t_j)$ is the discrete hazard component which is based on

$$H_0(t_j) = \sum_{j=1}^{k} \Delta H_0(t_j) \tag{7}$$

(Kalbfleisch and Prentice 2002, 114f.). Using this calculation, the survivor function can be approximated by the exponentiated, negative sum of estimated hazards until failure time t_j

$$S(t_j|x_i) = e^{-\sum_{j=1}^{k} \Delta H(t_j|x_i)} \tag{8}$$

(Ruhe 2016).[10]

To demonstrate that this approximation of the survivor function yields good estimates of the true survivor function even with a limited sample size, I conduct a Monte Carlo simulation.[11] Thereby, I simulate data for differently shaped baseline hazard rates and various parameter specifications. The simulated data generating process comes from a Weibull model with increasing, decreasing as well as flat baseline hazards, i.e., with shape parameters $p = 0.75$, $p = 1$ as well as $p = 1.25$. The model includes a time-varying effect of a binary predictor variable x, whereby an observation has $x = 1$ when a random draw from a standard normal distribution

---

9   This framework also allows to incorporate time-varying covariates, in which case the covariate vector $x_{i(t)}$ depends on time. The supplementary material provides a hypothetical example.
10 See Putter *et al.* (2005) for a similar calculation using the Breslow estimator.
11 The replication material is available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/4J48AX.

returns a positive number. I show results for four different types of time-varying effects, based on the following data generating processes:[12]

A negative hazard ratio, which turns positive

$$h(t|x_i) = pt^{p-1}e^{\ln(0.05)-0.8x_i+0.3\ln(t)x_i}. \tag{9}$$

A positive hazard ratio, which turns negative

$$h(t|x_i) = pt^{p-1}e^{\ln(0.05)+0.8x_i-0.3\ln(t)x_i}. \tag{10}$$

A positive hazard ratio, which increases in size

$$h(t|x_i) = pt^{p-1}e^{\ln(0.05)+0.8x_i+0.3\ln(t)x_i}. \tag{11}$$

A negative hazard ratio, which increases in size

$$h(t|x_i) = pt^{p-1}e^{\ln(0.05)-0.8x_i-0.3\ln(t)x_i}. \tag{12}$$

The corresponding, analytically derived survivor functions for each data generating process are documented in the supporting material.

For each data generating process, I generate 100 data sets with 500 failure times each. For each data set I estimate the survivor function for $x = 1$ using Equation (8). To quantify the prediction error of the approach, I calculate the difference between the estimated survivor function and the true, analytically derived survivor function. Figure 4 plots the distribution of this prediction error over time for each data generating process. The solid line gives the estimated average error based on a local polynomial smoother. The dashed lines document the median as well as 5th and 90th percentile for the prediction error in bins with width of one analysis time unit. Figure 4 indicates that the median and the estimated mean are virtually identical and always close to zero, suggesting no systematic bias. At the same time, the variance of the prediction error is symmetric and quite small, as about 90 percent of the estimates display an error of 5 percentage points or less. The supporting material includes similar graphs for simulations with smaller sample sizes ($N = 200$ and $N = 50$). The replication material also provides Stata code on how to implement the calculations in Equation (8) using the user-written package described in Ruhe (2016). A tutorial for R as well as SAS is provided by Thomas and Reyes (2014). In the next section, I demonstrate how survivor functions substantively improve the interpretation of time-varying effects.

6 Empirical Example: The Time-varying Effect of Mediation

I highlight the intricate interpretation of time-varying effects with the important example of how effectively international third-party mediation appeases armed conflict.[13] Prominent research suggests that mediators may only have a short-term effect (cf. Beardsley 2008, 2011; Quinn et al. 2013) and that third-party pressure is correlated with shorter peace (Werner and Yuen 2005). Beardsley (2008) reconciles positive and negative conclusions about mediation effectiveness in the literature based on his finding that the risk of renewed conflict is initially lower if a mediator was involved. However, mediated cases are exposed to a higher risk of crisis recurrence after several years. These results are interpreted as an indication that mediators might face a dilemma of buying short-term peace at the expense of long-term stability (Beardsley 2008, 2011).

---

12 The data is generating using the approach provided by Crowther and Lambert (2012).
13 The replication material is available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/4J48AX.
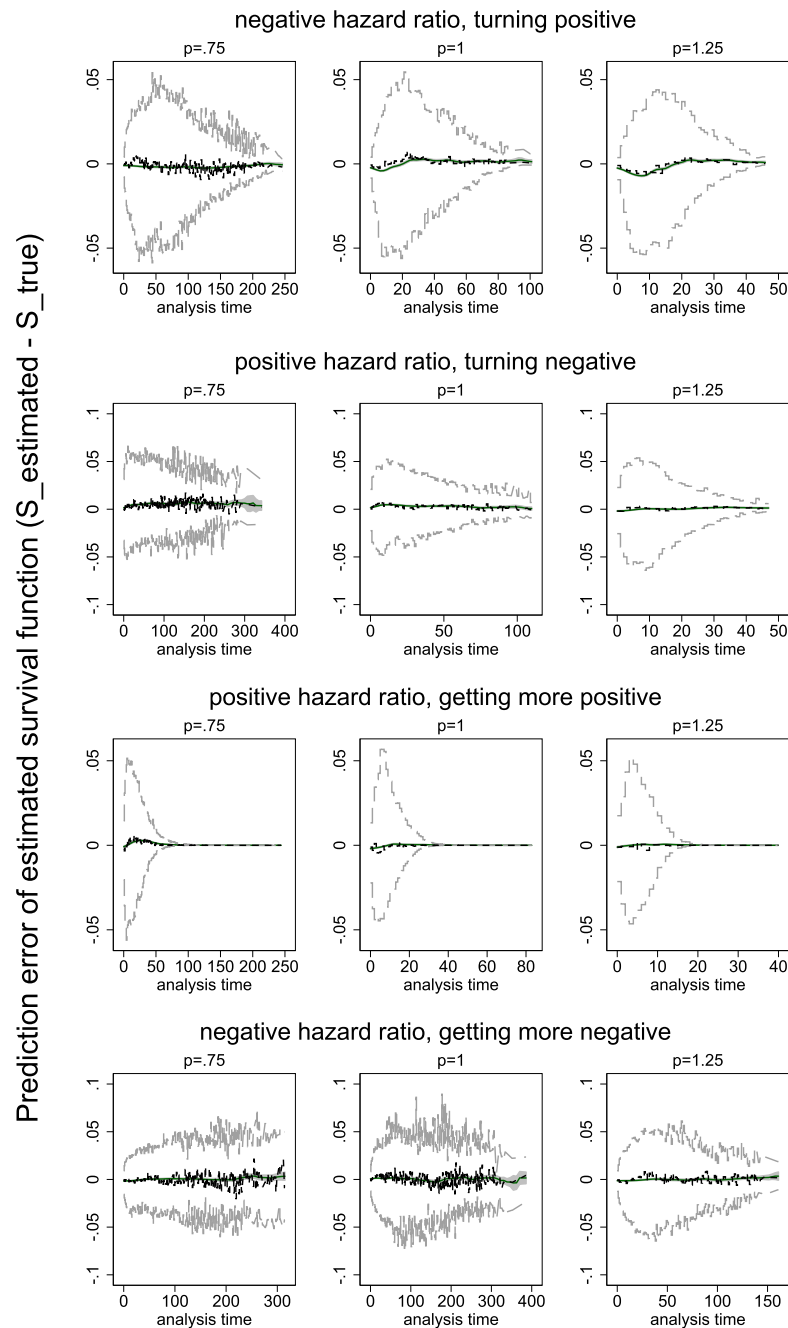
Figure 4. Monte Carlo Experiments: Distribution of the prediction error for data generating processes (9)–(12). Solid line gives estimated average based on local polynomial smoother. Dashed lines give 5th, 50th and 90th percentile of the error calculated in bins (width=1 analysis time unit).

If this interpretation were true, diplomats trying to appease international conflicts would face a difficult trade-off. It would beg the question of at what point and under which conditions the short-term benefits of mediation are outweighed by its long-term problems. Furthermore, if mediators were in fact buying short-term peace at the expense of long-term stability, would it be recommendable to get involved in the first place? These are all questions about the cumulative impact of mediation which cannot be answered with existing methods. In the following section, I show that, at scrutiny, the implications are less dramatic than the original interpretation of the time-varying effect might imply. Survivor functions are crucial to arrive at a less ambiguous conclusion.

The empirical analysis of the "mediation dilemma" is based on survival analyses of the duration of postconflict peace in which the effect of mediation is modeled through an interaction with time. The Cox model used by Beardsley (2008, 2011) therefore implies the following specification:

$$h(t|x_i) = h_0(t)e^{x_i\beta + \text{mediation}_i(\delta_1 + \delta_2 t)}, \tag{13}$$

whereby the effect of the binary predictor mediation is allowed to vary with a linear function of time. The empirical analysis indeed finds a strongly significant coefficient for the interaction with time and indicates that the hazard rates of mediated and unmediated cases cross after a few years. This indicates an effect of type (b) in Figure 1. Beardsley (2008, 737) concludes from this finding that "crisis dyads with mediation [...] are less likely to experience a recurrence of crisis within the first few years after a crisis. Yet mediators tend to only produce a pause before the dyad eventually becomes even more prone to recurrence than if it had not had mediation".[14]

While the conclusion is accurate about the instantaneous effect, Beardsley also draws conclusions about the cumulative effect. Based on the crossing hazard rate and the sign change in the hazard ratio, Beardsley states that despite positive short-term effects "[i]n the long run, mediation can create artificial incentives that, as the mediator's influence wanes and the combatants' demands change, leave the actors with an agreement less durable than one that would have been achieved without mediation" (Beardsley 2008, 723). Beardsley (2011) elaborates on this potential dilemma in much greater detail.

This strong claim is unfortunate, since it suggests to scholars and practitioners that mediation worsens the durability of peace in the long-term. Thus, it makes a statement about the cumulative effect of mediation over time, which cannot be drawn reliably from the hazard rates and hazard ratios presented in the paper. Nevertheless, the interpretation of Beardsley's results could still be correct. Time-varying hazard ratios and crossing hazard rates are simply too ambiguous. With only these tools available at the time of the study, a more informed conclusion was not directly available. Under these circumstances, it was important that Beardsley highlighted this possibility. Due to the immense policy relevance, however, it is of great importance to understand the ambiguity inherent in the results and to reanalyze the question with appropriate methods.

As described before, the interpretation of the empirical hazard rate estimates is not sufficient to make a claim on the cumulative effect. In fact, such an interpretation misses the fact that a hazard rate represents the *instantaneous* probability of failure, *conditional* that an event has not occurred up to this point in time (cf. Box-Steffensmeier and Jones 2004). It appears, that Beardsley (2008, 737) alludes to this important fact in a small paragraph in his conclusion: "Moreover, the results should not be interpreted as suggesting that mediated crises are *unconditionally* more likely to recur. Recall that unmediated peace arrangements are much more likely to fail in the first few years after a crisis. The key point is that mediation does very well in sustaining short-term peace at the expense of some potential for extremely durable peace. [Emphasis in the original]" Unfortunately, however, this important detail is overlooked for the remaining parts of the paper and most parts of the book. Immediately after the statement above, Beardsley (2008, 737) again evokes that mediation creates a trade-off between short- and long-term effects.

To understand under which circumstances mediation would be associated with long-term problems, we can use the analogy that mediation acts like a medical drug against renewed conflict. The "treatment" mediation is intended to stabilize the "immune system" of the most

---

14 Beardsley (2008, 2011) also uses a discrete duration model with a probit link function and cubic polynomials. This alternative form of a duration model yields very similar information as a Cox model by modeling the effect of variables on the discrete hazard probability, not a continuous hazard rate (see Box-Steffensmeier and Jones 2004, 71f.). Since the reanalysis of time-varying effects in the discrete models leads to the same substantive conclusions, I focus on Beardsley's Cox model below.

unstable conflict dyads. Over time, the concentration of the drug in the body decreases. For example, the mediator might become less involved, stop monitoring or the agreement fostered by the mediator is no longer adequate due to changing conflict parties. Hence, the effect of mediation diminishes with time and, at some point, exerts no influence anymore. In the drug example, the concentration of the drug in the body has been reduced to zero. If we want to know whether mediation actually creates long-term problems which make mediated cases worse off, the medical analogy helps to highlight what kind of pattern we are looking for: We suspect that the "treatment" mediation creates adverse effects or side effects, rather than just losing all influence. Unfortunately, however, crossing hazard rates as shown in panel (c) of Figure 2 as well as Figure 3 in Beardsley (2008, 736) and Figure 5.1 in Beardsley (2011, 113) are no indication of adverse effects.

As discussed above, survivor functions are needed to investigate whether the overall long-term effect of mediation is problematic. Hence, I use the data from Beardsley (2011) to replicate the original result and calculate the survivor functions implied by the model. I begin with the initial, bivariate comparison and plot the hazard rate of a renewed crisis for both mediated and unmediated crisis dyads. This corresponds to the analysis reported in Figure 3 in Beardsley (2008, 736) and Figure 5.1 in Beardsley (2011, 113). To replicate the results with as little assumptions as possible I rely on a fully nonparametric analysis. I estimate the smoothed hazard estimates as well as Kaplan–Meier survival functions for mediated and unmediated crisis outcomes using the data from (Beardsley 2011). The results in Panel (a) of Figure 5 bear a striking resemblance with the results reported in Beardsley (2008, 2011).[15] Since the analysis time unit is a single day, the values of the hazard rate denoted on the y-axis are much smaller. However, the overall pattern is very similar to the original results. Panel (b) in Figure 5 plots the corresponding Kaplan–Meier survival estimates. It becomes apparent that in a purely descriptive situation without control variables, there is always a higher proportion of 'surviving' cases in the mediated than in the unmediated dyads. This implies that mediated cases are more likely to remain at peace throughout the study period. Toward the end, however, the difference becomes quite small and no longer statistically significant.

Substantively, the descriptive evidence does not support a mediation dilemma. Rather, for quite some time, mediated cases are somewhat more stable than unmediated cases. Consequently, based on this evidence alone, mediators do not seem to face a trade-off between achieving short-term success at the expense of long-term stability. However, it has to be kept in mind that these results might suffer from considerable confounding. I therefore replicate Beardsley's core model documented in chapter 5, which is substantively identical to the central model in Beardsley (2008). The Cox model estimates the effect of covariates on the duration until a renewed crisis breaks out. The observations are censored after 10 years, or 3650 days. Since the data consists of multiple failure time data, the model stratifies for the number of previous crises

---

15 In fact, panel (a) is the nonparametric and (daily) continuous-time equivalent of the dyad-year discrete duration model without control variables used by Beardsley. The graph in Beardsley (2008, 736) as well as Beardsley (2011, 113) displays the hazard probability $h(t)$ from the following discrete duration model

$$h(t \mid \text{mediation}_i) = P(T = t \mid T \geq t, \text{mediation}_i)$$
$$= \Phi(\beta_0 + (\beta_1 + \beta_2 t)\text{mediation}_i + \beta_3 t + \beta_4 t^2 + \beta_5 t^3), \qquad (14)$$

whereby Φ is the standard normal cumulative distribution function and the baseline hazard probability is modeled using cubic polynomials of time (cf. Box-Steffensmeier and Jones 2004, 71ff.). This is the discrete time equivalent to the hazard rate, i.e., the probability of failure at time t, given survival until time t.

Based on the predicted hazard probability $h(t)$ we can calculate the survival function for Beardsley's discrete duration model as

$$S(t \mid \text{mediation}_i) = P(T \geq t \mid \text{mediation}_i) = \prod_{j=1}^{t}(1 - h(t_j \mid \text{mediation}_i)) \qquad (15)$$

(cf. Box-Steffensmeier and Jones 2004, 72).

I show in the supplementary material that replicating Beardsley's discrete model with control variables and calculating survival functions produces substantively identical results as Figure 5(b).
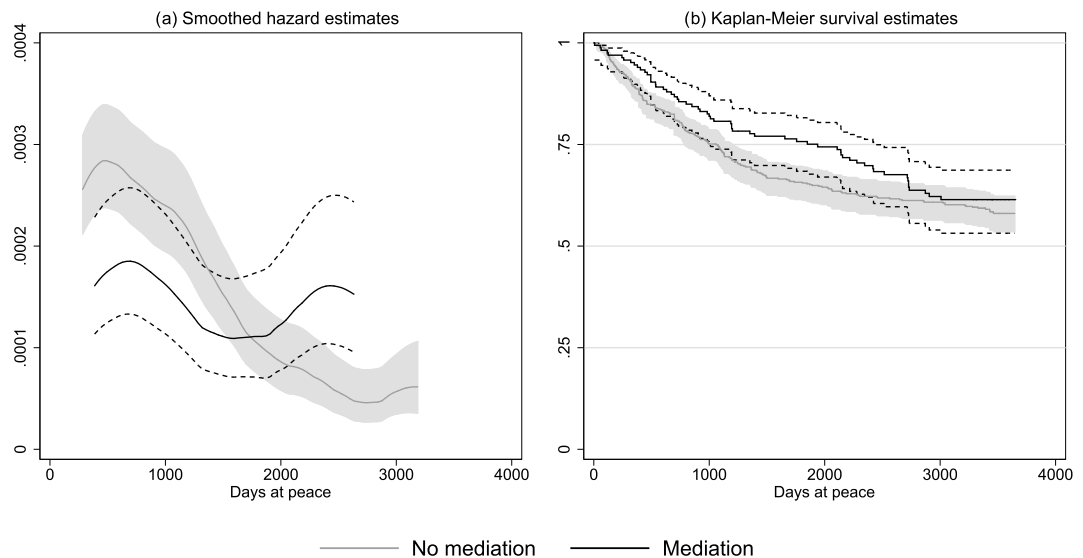
Figure 5. Comparing the stability of mediated and unmediated agreements: Smoothed hazard and Kaplan–Meier survival estimates using the data from Beardsley (2011, 208ff.).

which a dyad experienced. The model uses several predictor variables, all of which are interacted with a linear function of time. These predictor variables capture the number of previous crises, the violence level of the conflict, the natural log of the crisis duration, and dichotomous indicators if both sides in the dyad are a democracy, whether the conflict ended in victory as well as whether the states are territorially contiguous (Beardsley 2011, 208ff.). The complete regression model by Beardsley (2011) interacts all predictor variables with a linear function of time.

Table 1 provides the estimates of the analysis. Model 1 is an exact replication of Beardsley's model (2011, app. c. 5). Most of the interactions are highly statistically significant. However, at closer inspection, Model 1 still violates the proportional hazards assumption for virtually every single variable according to a test using Schoenfeld residuals. Keele (2010) describes how an incorrectly specified model may lead to a significant test statistic. Moreover, including interactions with time when the proportional hazards assumption is not violated may create such a violation based on a misspecified model (Box-Steffensmeier and Jones 2004, 136, n. 8). Hence, I modify the model. It appears that a model in which only the mediation effect is allowed to vary with time fits the data generating process best and already fulfills the proportional hazard assumption. Model 2 in Table 1 documents the coefficient estimates for this restricted model specification.

Regarding the effect of mediation, both models appear to estimate a substantively similar pattern. Mediation originally decreases the hazard of a renewed crisis. Eventually, however, this effect is reversed. As described above, this pattern in itself does not imply that mediation has a counterproductive long-term effect. In order for this to be the case, the survival curves for mediated and unmediated cases would need to cross after a certain time. The simple bivariate comparison without control variables in Figure 5 suggests that this is not the case. However, the effect of the control variables may alter this conclusion.

Figure 6 depicts for both the original as well as the revised, restricted model the estimated survival function for mediated and unmediated cases if all remaining variables are held at their mean value in the sample. Since different shapes of the baseline hazard rate may affect the magnitude of mediation's time-varying effect on crisis risk, the plot distinguishes between the five strata in the model.[16] The results clearly indicate that for the average case, mediation is in no

---

16 In a model with multiple time-varying effects it would be important to plot results for different, meaningful values of these variables, to assess if the substantive conclusions differ across scenarios.

Table 1. Mediation and crisis recurrence: Replication of the core model in (Beardsley 2011, 208ff., table 5.1, model 1).

	(1) Original Cox	(2) Restricted Cox
Mediation	−0.913**	−0.681**
	(0.264)	(0.218)
Mediation × time	0.000539*	0.000368**
	(0.000212)	(0.000138)
Previous crises	0.419**	0.0582
	(0.127)	(0.063)
Violence level	0.232	−0.178**
	(0.136)	(0.0563)
Crisis duration	0.302**	−0.0831*
	(0.0857)	(0.0399)
Democratic dyad	−0.605	−0.784*
	(0.626)	(0.31)
Victory	0.607**	−0.0212
	(0.209)	(0.123)
Contiguity	1.24**	0.559**
	(0.342)	(0.148)
Previous crises × time	−0.000328**	
	(0.000105)	
Violence level × time	−0.000368**	
	(0.000102)	
Crisis duration × time	−0.000337**	
	(0.0000632)	
Democratic dyad × time	−0.0000461	
	(0.000406)	
Victory × time	−0.000543**	
	(0.000152)	
Contiguity × time	−0.000589**	
	(0.000197)	
Cases	705	705
Failures	276	276

Note: cluster robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$.

way associated with long-term problems. On the contrary, mediated settlements are estimated to remain substantively more stable than unmediated crisis outcomes, despite the fact that the advantage of mediated crises eventually decreases. Figure 6 further shows that this result holds across strata and regardless of whether the original or the restricted model is used. Due to the uncertainty in the parameter estimates and the survival estimates, the difference in survival probability becomes insignificant after five to eight years.[17] This confirms the Kaplan–Meier estimates in Figure 5.

---

17 See supplementary material.
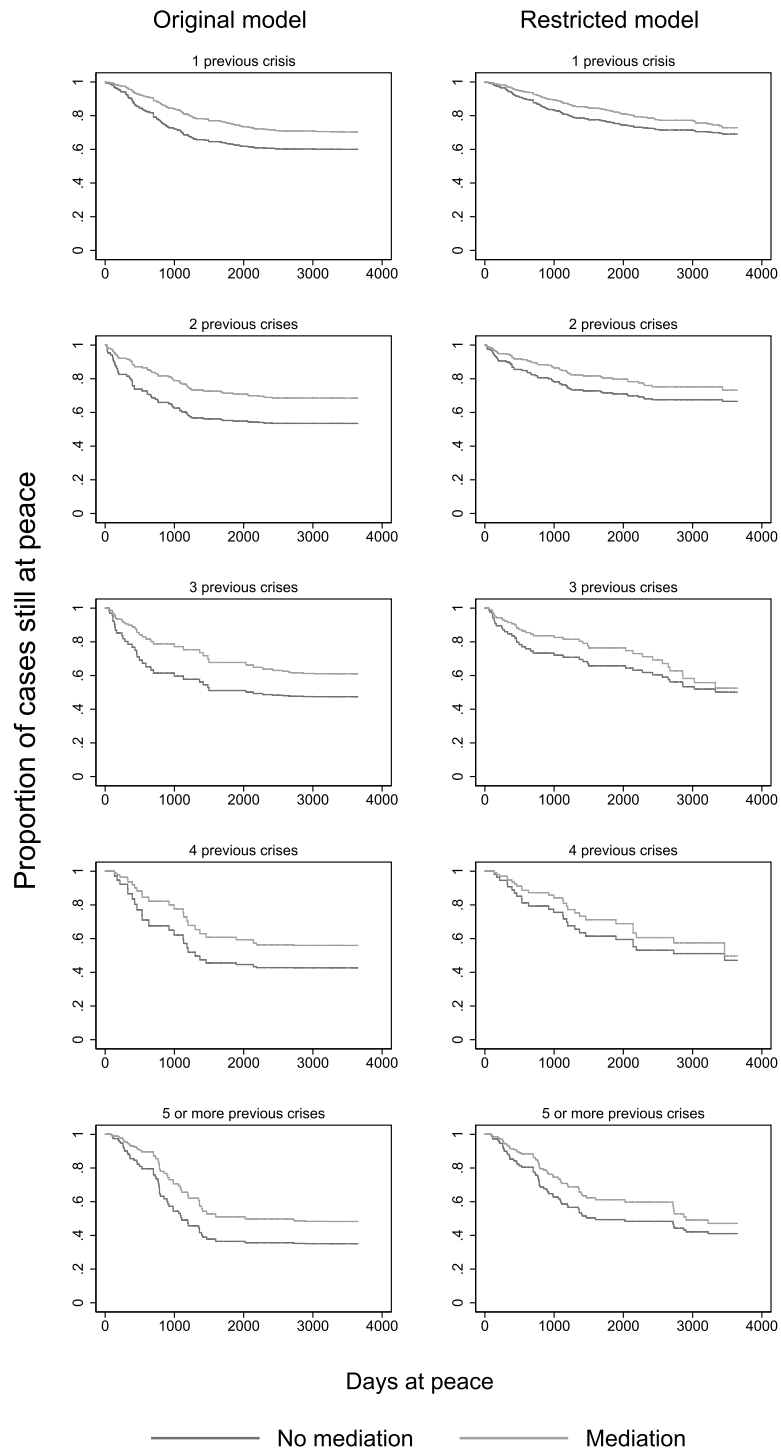
Figure 6. Predicted effect of mediation on the duration of postcrisis peace. Results are reported for each stratum based on the estimates from all models in Table 1. All variables, except mediation, held at their mean.

7 Discussion and Recommendations

The analysis of the time-varying mediation effect provides a more nuanced image of mediation effectiveness than earlier studies. The replication confirms empirical evidence which shows that, compared to unmediated crises, mediated cases are more stable early on in a postconflict period, but less stable at later points in time, given that they did not yet experience a renewed crisis and compared to unmediated cases which also did not yet fail. However, the empirical evidence

provides no indication that fewer mediated cases are at peace after several years than unmediated cases. Hence, the results do not support the hypothesis of a mediation dilemma.

The results from the reanalysis of Beardsley (2011) suggest that mediation is associated with substantively more stable conflict outcomes, although this difference becomes small and statistically insignificant after several years. This implies that while both mediated and unmediated dyads might eventually relapse into crisis, mediated cases do so later. In other words, mediation is associated with peace for some time, but not indefinitely. How these results compare across different mediators of alternative mediation strategies should be analyzed in further analyses. This paper introduces the necessary methodology in the form of survival functions.

On a larger scale, the example and the discussion of time-varying effects clearly indicate that neither hazard ratios nor hazard rates for specific scenarios provide the full picture of the substantive meaning of time-varying effects with a change in sign. Depending on the baseline, the cumulative impact of such time-varying effects can be both a drastic reversal of an effect as well as no substantive change at all. This shows that an interpretation based on time-varying hazard ratios or hazard rates alone leave a lot of ambiguity regarding the overall effect, since it is not clear if mediated cases are on average worse off in the long run. However, appropriate visualizations using survivor functions are able to reduce this ambiguity and provide a clearer picture of how the overall effect evolves over time.

This leads to the following recommendations for researchers dealing with time-varying effects in duration analyses. These recommendations consist of four steps and extend earlier research on the interpretation of nonproportional hazards:

1. In order to identify and model nonproportional hazards, the steps outlined by Keele (2010) as well as Park and Hendry (2015) should be followed.
2. Time-varying effects can thereafter be analyzed using hazard ratios or relative hazards as described by Licht (2011). This allows visualizing the pattern of the instantaneous, multiplicative time-varying effect and helps to assess if the effect significantly changes its sign (e.g., pattern b in Figure 1).
3. If the effect significantly changes its sign, it is recommendable to clarify the substantive cumulative effects using survivor functions as outlined above. If the effect does not change its sign, survivor functions may nevertheless provide an intuitive summary of the estimated cumulative effects and the pattern in the data.
4. Researchers should be aware that a variable's effect on the survival function could vary, depending on the baseline hazards across different strata as well as due to the values of other (time-varying) covariates in the model. Thus, survivor functions should be used to intuitively communicate predictions for different, meaningful scenarios.[18]

8 Conclusion

This paper demonstrates that modeling violations of the proportional hazards assumption using interactions with time makes a correct interpretation of covariate effects very complex. Neither time-varying hazard ratios nor hazard rates for specific covariate values are sufficient to describe the overall substantive effect and are very ambiguous if a time-varying effect changes its sign. The presence of multiple time-varying effects further complicates the inference, since the shape of the baseline varies with the values of the covariates. To describe a variables overall effect, researchers should use survivor functions for meaningful covariate values in order to enable an intuitive and unambiguous inference.

---

18 As described in Equation (6), time-varying covariates can be included if these changes are an important aspect of the underlying process, e.g., when a change in a variable of interest occurs only after some time. However, researchers need to consider the assumptions associated with time-varying covariates in a duration model (see Box-Steffensmeier and Jones 2004, 95ff.).

Using these statistics, the reanalysis of mediation effectiveness in interstate conflicts provides a more optimistic conclusion than earlier research. The visualization with survivor functions shows that the average mediator *does not* create short-term peace at the expense of long-term stability. Hence, mediation *does not* entail a potential trade-off between short- and long-term stability. Instead, the findings suggest a much more encouraging policy implication: Mediated agreements appear to be considerably more stable than unmediated conflict outcomes, before they eventually converge to a similar stability level as in unmediated conflicts.

Supplementary materials

For supplementary materials accompanying this paper, please visit
https://doi.org/10.1017/pan.2017.35.

References

Allen, S. H. 2005. The determinants of economic sanctions success and failure. *International Interactions* 31(2):117–138.

Beardsley, K. 2008. Agreement without peace? International mediation and time inconsistency problems. *American Journal of Political Science* 52(4):723–740.

Beardsley, K. 2011. The mediation dilemma. In *Cornell studies in security affairs*. Ithaca, NY: Cornell University Press.

Beck, N. 2010. Time is not a theoretical variable. *Political Analysis* 18(3):293–294.

Beck, N., J. N. Katz, and R. Tucker. 1998. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42(4):1260–1288.

Box-Steffensmeier, J. M., and B. S. Jones. 2004. *Event history modeling: A guide for social scientists. Analytical methods for social research*. New York: Cambridge University Press.

Box-Steffensmeier, J. M., D. Reiter, and C. J. W. Zorn. 2003. Nonproportional hazards and event history analysis in international relations. *Journal of Conflict Resolution* 47(1):33–53.

Box-Steffensmeier, J. M., and C. J. W. Zorn. 2001. Duration models and proportional hazards in political science. *American Journal of Political Science* 45(4):972–988.

Brambor, T., W. R. Clark, and M. Golder. 2005. Understanding interaction models: Improving empirical analyses. *Political Analysis* 14(1):63–82.

Carter, D. B., and C. S. Signorino. 2010a. Back to the future: Modeling time dependence in binary data. *Political Analysis* 18(3):271–292.

Carter, D. B., and C. S. Signorino. 2010b. Reply to time is not a theoretical variable. *Political Analysis* 18(3):295–296.

Chiozza, G., and H. E. Goemans. 2004. International conflict and the tenure of leaders: Is war still ex post inefficient? *American Journal of Political Science* 48(3):604–619.

Cleves, M. A., W. Gould, R. G. Gutierrez, and Y. V. Marchenko. 2010. *An introduction to survival analysis using Stata*. 3 ed. College Station, TX: Stata Press.

Cox, D. R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society Series B - Statistical Methodology* 34(2):187–220.

Crowther, M. J., and P. C. Lambert. 2012. Simulating complex survival data. *Stata Journal* 12(4):674–687, (14).

Gandrud, C. 2015. simph: An R package for illustrating estimates from cox proportional hazard models including for interactive and nonlinear effects. *Journal of Statistical Software* 65(3):1–20.

Golub, J. 2007. Survival analysis and European Union decision-making. *European Union Politics* 8(2):155–179.

Golub, J., and B. Steunenberg. 2007. How time affects EU decision-making. *European Union Politics* 8(4):555–566.

Grewal, S., and E. Voeten. 2015. Are new democracies better human rights compliers? *International Organization* 69(02):497–518.

Hale, T. 2015. The rule of law in the global economy: Explaining intergovernmental backing for private commercial tribunals. *European Journal of International Relations* 21(3):483–512.

Kalbfleisch, J. D., and R. L. Prentice. 2002. *The statistical analysis of failure time data*. 2 ed. Hoboken, NJ: Wiley-Interscience.

Keele, L. 2010. Proportionally difficult: Testing for nonproportional hazards in Cox models. *Political Analysis* 18(2):189–205.

Licht, A. A. 2011. Change comes with time: Substantive interpretation of nonproportional hazards in event history analysis. *Political Analysis* 19(2):227–243.

Murillo, M. V., and C. Martínez-Gallardo. 2007. Political competition and policy adoption: Market reforms in Latin American public utilities. *American Journal of Political Science* 51(1):120–139.

Park, S., and D. J. Hendry. 2015. Reassessing schoenfeld residual tests of proportional hazards in political science event history analyses. *American Journal of Political Science* 59(4):1072–1087.

Putter, H., M. Sasako, H. H. Hartgrink, C. J. H. van de Velde, and J. C. van Houwelingen. 2005. Long-term survival with non-proportional hazards: Results from the Dutch gastric cancer trial. *Statistics in Medicine* 24(18):2807–2821.

Quinn, D. M., J. Wilkenfeld, P. Eralp, V. Asal, and T. Mclauchlin. 2013. Crisis managers but not conflict resolvers: Mediating ethnic intrastate conflict in Africa. *Conflict Management and Peace Science* 30(4):387–406.

Royston, P., and M. K. B. Parmar. 2011. The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in Medicine* 30(19):2409–2421.

Ruhe, C. 2017. Replication data for: Quantifying change over time: Interpreting time-varying effects in duration analyses. Harvard Dataverse, doi:10.7910/DVN/4J48AX, V1, UNF:6:NqnTUN7yBqlg7xJpgdfAmw==.

Ruhe, C. 2016. Estimating survival functions after stcox with time-varying coefficients. *Stata Journal* 16(4):867–879.

Thomas, L., and E. M. Reyes. 2014. Tutorial: Survival estimation for Cox regression models with time-varying coefficients using SAS and R. *Journal of Statistical Software* 61(1):1–23.

Werner, S., and A. Yuen. 2005. Making and keeping peace. *International Organization* 59:261–292.

Williams, L. K. 2016. Long-term effects in models with temporal dependence. *Political Analysis* 24(2):243–262.

Zhelyazkova, A., and R. Torenvlied. 2009. The time-dependent effect of conflict in the council on delays in the transposition of EU directives. *European Union Politics* 10(1):35–62.