Paid Medical Malpractice Claims: How Strongly Does the Past Predict the Future?

Kowsar Yousefi
University of Tehran, Economics Department

Bernard Black
Northwestern University, Pritzker School of Law and Kellogg School of Management

David A. Hyman
Georgetown University School of Law

(draft July, 2022)

Paid Medical Malpractice Claims: How Strongly Does the Past Predict the Future?

Kowsar Yousefi, Bernard Black & David A. Hyman

Abstract. Using hazard analysis, we study whether various physician characteristics, including prior paid claim history, gender, specialty, years of experience, type of degree (M.D. versus D.O.), country of medical school attendance (U.S. versus non-U.S.), and gender) predict future paid medical malpractice ("med mal") claims, using detailed data on all licensed physicians and all paid claims in Illinois over a 25-year period. This level of granularity is not available using national data. After controlling for other factors, physicians with a single prior paid claim have a four-fold higher risk of future claims than physicians with zero prior paid claims. Male gender, attending a non-U.S. medical school, and practicing in a high-malpractice-risk specialty all predict higher paid claim risk. Paid claim risk is also higher for physicians with 6-15 prior years of experience than for those who are either earlier or later in their careers. We find having an M.D. (rather than a D.O.) is associated with higher paid claim risk, but only in our multiple-failure models.

Keywords: Medical Malpractice, Survival Analysis

1. Introduction

There is a substantial literature on factors that predict a physician's risk of having a paid medical malpractice ("med mal") claim. Age, gender, specialty, and a history of past paid claims have all been found to predict claim risk. We contribute to this literature by using a novel dataset, which contains detailed personal information (including claiming history) for the 84,114 physicians with an active license in Illinois at any time over a 25-year period (1990-2014). We apply survival analysis to estimate the relative impact of multiple factors on claim risk, including the number and recency of prior paid claims, specialty, years of practice experience, whether the physician has an M.D. or D.O. degree, gender, and whether the physician attended a U.S. or non-U.S. medical school.

Prior studies of the predictive power of prior paid claims have focused on the minority of physicians – in many specialties, a small minority -- who have paid claims. Because we also have data on the 92% of physicians in Illinois with zero paid claims, we have a much richer dataset with which to analyze the factors that predict paid claim risk. We find that physicians with 1 prior paid claim have an annual risk of future paid claims that is roughly 4-fold higher than physicians with zero prior claims. The increase in future claim risk is 6-fold higher for physicians with 2 prior paid claims, and 8-fold higher for physicians with 3+ prior paid claims.

We also study the predictive effect of other physician characteristics. Overall, the hazard rate for female physicians is about half that of male physicians. Physicians with an M.D. degree have moderately higher paid claim risk than those with a D.O. degree. Board-certified physicians have a higher hazard rate than non-board-certified physicians. The hazard rate as a function of physician experience has an inverse U-shape; it is highest for physicians who have been practicing for 6-15 years, who are roughly 3 times more likely to have a claim than physicians in their first five years of practice, but declines with additional experience years. Physician who attended a non-U.S. medical school have higher paid claim rates than graduates of U.S. medical schools, but the predictive effect of medical school is limited to physicians without prior paid claims. Unsurprisingly, paid claim risk is higher in certain specialties (i.e., surgery, ob-gyn), but our other results are similar across both high-claim-risk and low-claim-risk specialties.

We present results principally using either the nonparametric Kaplan-Meier survival model or the Cox semiparametric proportional hazard model. However, we find very similar results using Gompertz and Weibull parametric models.

2. Literature Review

Most studies of med mal claim risk are retrospective, analyzing the association between claiming patterns and prior claims, as well as the impact of various demographic covariates (e.g., specialty, gender, years of experience, etc.)

Predictive value of prior claims. Bovbjerg and Petronis (1994) use Florida data from 1975 to 1988. They define a baseline period (1975-1980), and a subsequent period (1981-1988). They report that less than 7% of physicians had any claim (paid or unpaid) during the subsequent period, but the odds ratio for physicians with 1+ claims in the baseline period to have a claim during the subsequent is 2.84. Having even a single unpaid baseline claim approximately doubled the odds of a subsequent-period claim.

Studdert et. al. (2016) use the National Practitioner Data Bank (NPDB), a national repository of paid claims against physicians. Using an Anderson-Gill survival model, they assess the extent to which physicians with two or more paid claims face higher risk of a future paid claim compared to physicians with one paid claim. They find that "the risk of recurrence increased with the number of previous paid claims." Compared to physicians with a single paid claim, physicians with three paid claims had three times the risk of a future claim.

Black et. al. (2019) also use the NPDB and report that having even a single paid claim in an initial 5-year baseline period roughly quadruples the risk of having a paid claim in the next five years. Both of these studies report that prior paid claims predict future paid claims. However, Studdert et al. do not directly study the 94% of physicians in their sample with no paid claims during their study period, and Black et al. study these physicians only indirectly, by assuming an unobserved mass of physicians with no paid claims during their baseline period. A core limitation of both studies is that the NPDB contains data only on paid claims and even for physicians with paid claims, the public version available of the NPDB that is to researchers provides limited information on physician characteristics.

Gender, Specialty, Degree Type (M.D. or D.O.), and U.S. vs. non-U.S. Medical School. Taragin et. al. (1992) use New Jersey insurance data from 1977 to 1988. They divide physicians into three groups (high-claim rate, medium-claim rate, and low-claim rate), based on the number of claims per years at risk. They find that male physicians are 3x more likely to be in the high-claim rate group. Psychiatry had the lowest rate of claims per year. Internal medicine had 2.5x the number of claims per year as psychiatry. Neurosurgery, orthopedic surgery, and ob-gyn had 7x-12x the number of claims per year as psychiatry. They found no evidence of an association between the type of medical degree (M.D. or D.O.) and where one

attended medical school (U.S. or non-U.S. medical school) with claim rate group. However, age was associated with claim rate per year; the highest risk was found in physicians aged 36-55, with a peak at around 40 years.

More recent studies report similar findings. Unwin et. al. (2015) conduct an international meta-analysis of gender differences and find that male physicians have a claim risk that is 1.74 times that for female physicians. Jena et. al. (2011) use data from a national medical malpractice insurer to evaluate claim risk by specialty. They find neurosurgery has the highest risk of a med mal claim (19.1%), followed by cardio-thoracic surgery (18.9%) and general surgery (15.3%), compared to an overall claim rate across all physicians of 7.4%. Using Swedish data, Pukk-Harenstam et. al. (2008) similarly report that surgical specialties have the highest claim risk. In Australia, Nash et al. (2009) conducted a survey of practicing physicians, and found that those in certain specialties (i.e., surgery and ob-gyn) are more likely to be "involved in legal matters."

There is mixed evidence on the impact attending a non-U.S. medical school on claim risk. Studdert et al. (2016) and Adamson et al. (1997) both find a higher claim risk for physicians who attend non-U.S. medical schools, while (as noted previously) Taragin et. al. 1992) do not.

Our study is similar to Studdert et. al. (2016) and Black et. al. (2019) in that we use dynamic modeling of the hazard risk as a function of prior claims and other covariates. However, unlike these studies, we have data on all licensed physicians over an extended time period, including physicians who did not have any claims. We also have more detailed physician covariates. However, our study is limited to a single state (Illinois).

3. Data and Methods

3.1. Illinois dataset

We describe our Illinois dataset in greater detail in Hyman, Rahmati & Black (2021). We study Illinois physicians over a 25 year period (1990-2014).[1] Our data includes detailed information on each physician, including the date they were first licensed in Illinois; the date they were last licensed; self-reported specialty; whether they attended a non-U.S. medical

[1] Our data runs through 2016, but we drop 2015 and 2016 because there is a large drop-off in claim rates in 2015 and 2016 which appears to reflect right censoring of the dataset for those years. There were 206 paid claims in our Illinois dataset for 2014, compared to 134 in 2015 and 56 in 2016. There is no similar drop in paid claims reported by Illinois physicians to the NPDB.

school; any disciplinary actions against them in Illinois; and payments on med mal claims. Our data does not include gender, but we impute gender based on physicians' first names.

For paid med mal claims, we use the reported "settlement date" as the date of payment.[2] We drop the small number of physicians licensed before 1945 (224 observations); these physicians would have been at least 70 years old at the start of our sample period. We also drop 41 physicians whose license is reported to have expired before it was issued; 96 claims that are reported to have occurred prior to license issuance;[3] and 18 duplicate claims. We are left with a dataset of 9,827 paid claims for 84,114 physicians licensed at some time during 1990-2014. Some physicians had multiple paid claims, while most (77,256 (92%)) had none. Most of the physicians in our dataset attended U.S. medical schools, but 8,691 (10.3%) physicians attended non-U.S. medical schools.

Physicians are not required to report their specialty, so only 55% do so.[4] As we describe in Hyman, Rahmati & Black (2021), we identified twenty-two specialties, of which we classify nine as high-claim risk based on Illinois paid claim rates. We define high-claim-risk as specialties where at least 10 percent of physicians in that specialty have one or more paid claims during our sample period. The high-claim-risk specialties are Obstetrics and gynecology, Otolaryngology/ENT, Urology, General Surgery, Orthopedic surgery, Cardio-thoracic surgery, Neurosurgery, Vascular surgery, and Colorectal surgery). Appendix Table A-1 lists all 22 specialties, indicates which are high-claim-risk, and provides information on claim risk by specialty. Our categorization of high-claim-risk specialties is consistent with U.S. evidence (Jena, 2011) and broadly consistent with Swedish evidence (Pukk-Harenstam et. al., 2008).

Because we know the date that each physician was first licensed in Illinois, we can analyze patterns of claiming by experience year (i.e., settlement year – year of initial licensure). Across our dataset, the mean (median) number of experience years is 18.4 (16.4). We impute gender based on first name. For 7,189 physicians (8.5%), principally with foreign names, we

[2] For 23 cases with missing settlement date, we instead use "court date" (which is not defined in the IDFPR dataset).

[3] One possible reason for the existence of these claims: The Illinois reporting requirement is not limited to paid claims resulting from practice in Illinois. It is possible that a few physicians, moving to Illinois from elsewhere, reported prior paid claims at the time they applied for an Illinois license, or had paid claims reported by their insurers that were attributable to their pre-Illinois practice.

[4] Physicians also self-report whether they are board certified, and of those who report a specialty, 92% report being board-certified.

were unable to impute gender; we drop these physicians in regressions which include gender as a covariate.

Table 1 and Appendix Tables A-3 and A-4 provide basic demographic information on the physicians in our data.

Table 1: Summary Information on Physician Demographics

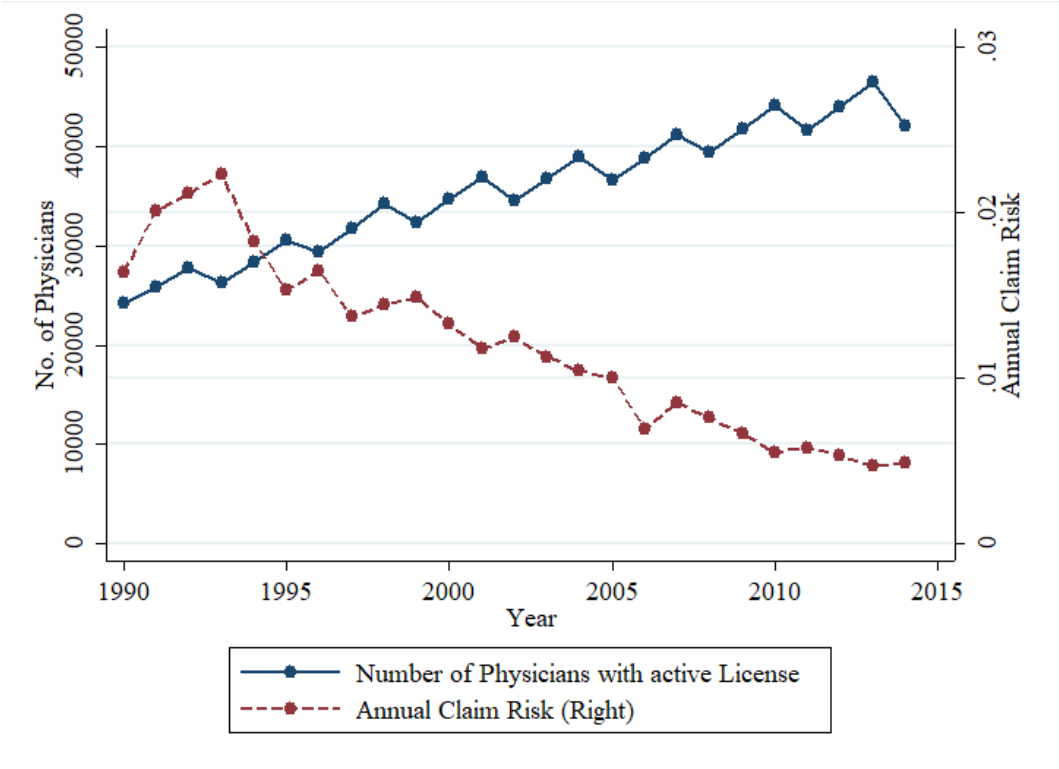
Summary information for 84,114 physicians, licensed in Illinois at any time over 1990-2014. Gender is imputed from first name. Specialty, medical school attended, and medical degree are self-reported. High-claim-risk specialties are defined as specialties where at least 10 percent of physicians in that specialty have one or more paid claims, per Hyman, Rahmati & Black (2021). See Table A-1 for a list of high-claim-risk specialties.

Physician Demographics		No. of Physicians	Share,%
Gender	Male	53,424	63.5
	Female	23,501	27.9
	Uncertain	7,189	8.55
Board Certified?	Yes	42,472	50.05%
	No	3,602	4.28%
	Unknown	38,040	45.22%
Medical School Attended	U.S.	23,997	28.53
	Non-U.S.	8,691	10.33
	Unknown	51,426	61.14
Specialty	High-claim-risk	7,698	9.15
	Low-claim-risk	38,376	45.62
	Unknown specialty	38,040	45.22
Medical Degree	M.D.	67,153	79.84
	D.O.	5,280	6.28
	Unknown	11,681	13.89
No. of Paid Med-Mal Claims	Zero	77,256	91.84
	1	4,814	5.72
	2	1,290	1.53
	3+	754	0.89

Figure 1 shows time trends in the number of physicians with an active license, and annual claim risk. There is steady growth in the number of licensed physicians and, from a peak in 1993, a steady fall in the annual risk of a paid med mal claim. The drop in paid claim risk is consistent with national trends (Paik, Black, and Hyman, 2013).

Figure 1: Number of Active Licensed Physicians and Annual Claim Risk

Figure shows number of physicians with an active license (left-hand y-axis) and the annual claim risk (number of paid claims in each year divided by the number of active physicians in that year (right-hand y-axis)). See Table A-2 for numerical values for licensed physician and paid claims.



3.2. Survival Models

Our survival analysis relies on several different hazard models. The simplest is the Kaplan-Meier nonparametric model, which we use to compare survival rates for different groups. We use both single-failure and multiple-failure models. For the single- failure model, the Kaplan-Meier survival rate is obtained by counting, for each day, the number of physicians with a paid claim, divided by the number of surviving, licensed physicians on that day, taking into account attrition, and nonsurvival (having a paid claim). Attrition (treated in the model as right-censoring) occurs on the day when a physician's Illinois license expires; nonsurvival occurs on the date a claim is paid. The ratio of paid claims to surviving, licensed physicians on each day is then subtracted from the survival rate from the prior day. The starting point for the survival analysis is January 1, 1990; we limit the sample to physicians licensed over 1990-

1994 (including physicians first licensed during 1990, these physicians enter the sample for the survival model on the day they are licensed). Some results are further restricted to physicians licensed over 1990-2014.

We also use a multiple-failure model, known as an Anderson-Gill model. In this approach, entry after January 1, 1990 is allowed and occurs on the day when a physician gains an Illinois license. Also, a physician with a paid claim is added back to the pool of licensed physicians on the day after the paid claim, as if this physician had newly entered the population of licensed physicians. Multiple claims against a single physician are treated as independent.

We are also interested in the extent to which physician characteristics, including prior paid claims, gender, nature of medical degree, U.S. versus foreign medical degree, years of experience (we group physicians into experience groups; the groups we most often use are 0-5 years; 6-10, 11-15, and 16+ years of experience), and dummy variables for practicing in a high-claim-risk, low-claim-risk or unknown specialty (alternatively we use dummy variables for each specialty). We study the effect of these characteristics using both a univariate approach, based on the Kaplan-Meier model, and a multivariate approach, using the Cox semi-parametric survival model. For the Cox model results, we confirm in the Appendix that we find very similar results using the parametric Gompertz and Weibull models. We again use both a single-failure and a multiple-failure approach.

The Cox model is similar to the Kaplan-Meier model in how it handles entry into the sample, attrition from the sample, and non-survival (having a paid claim). It includes both a baseline survival function and a relative risk factor based on covariates. In this model, the proportion of physicians who survive to time *t* (do not face a paid claim up to that point) is given by:

$$\boldsymbol{Cox\ Model, survival\ form}\text{:}\ S_{it}(\boldsymbol{X}_i) = S_{0t}e^{X_i\beta} \qquad (1)$$

Here *i* indexes individual physicians; $\mathbf{X}_i$ is a vector of covariates and $\boldsymbol{\beta}$ represents the associated coefficients. The covariates (not all included in all models) are number of prior paid claims (alternatively, dummy variables for having 1, 2, or 3+ prior claims); years of experience (6-10 years; 11-15 years, 16+ years; 0-5 years is the omitted group); gender (male is omitted); medical degree (M.D., D.O. or unknown; D.O. is omitted); medical school (U.S., non-U.S. or unknown; U.S. is omitted). The baseline survival function (S_{0t}) gives the probability of being alive when all covariates are zero. The model is estimated using maximum likelihood. A core

assumption is that the covariates have independent effects on survival. The Cox model can also be specified in terms of a hazard rate per unit time:

$$\boldsymbol{Cox\ Model, hazard\ rate\ form}\text{:}\ h_{it}(\boldsymbol{X}_i) = h_{0t}e^{X_i\beta} \qquad (1)$$

The Weibull Gompertz and Weibull models are similar, but impose a specific functional form for the baseline hazard rate $h_0(t)$: Weibull ($h_0(t) = pt^{p-1}$)), Gompertz ($h_0(t) =$ exp (γt)).

4. Initial Analysis of the Predictive Power of Past Claims

In Table 2, we study the predictive power of the number of prior period claims for the likelihood of future claims, without controlling for other physician characteristics. We limit the sample to a balanced panel of the 15,881 physicians licensed throughout the full sample period (1990-2014). Consider first the row for the 14,440 physicians (90.9%) with no paid claims during 1990-1994. In the first set of columns, we report the actual number of physicians with no claims during the next 20 years (1995-2014), the expected number if paid claims arrived at random over this 20-year period, and the ratio of actual to expected number of physicians with 0 future-period claims. To simulate random claim arrival, we first count the actual paid claims for all 15,881 physicians over 1995-2014; this is 4,694 claims. We assume that these claims arrive at the same rate in each year (the actual claim rate falls over this period); thus the annual per-physician risk is (1/20)*(4,694/15,881) = 0.01478. We then assign each paid claim at random to a physician-year in the full sample.[5] Most physicians will have no randomly assigned claims, but some will have one assigned claim, a smaller number will have two assigned claims, a still smaller number will have three assigned claims, and so on. Each physician will thus have both an actual number of paid claims during the future period and an expected number of claims, if claims arrived at random. Table 2 shows actual vs. expected future paid claims (with columns for 0, 1+ and 2+ future paid claims), broken out separately for physicians with (0, 1, 2, and 3+ paid claims during the prior period (1990-1994).

[5] For each physician year in the future period, we draw a random number from the uniform distribution over [0, 1]. We assign an expected claim to that physician-year if the randomly drawn number is less than 0.01478.

Table 2. Future Claims: Actual vs. Expected (with Random Claim Arrival)

Comparison of actual number of paid claims over 20-year "future" period (1995-2014) for the 15,881 physicians, licensed throughout 1990-2014, with different numbers of paid med-mal claims during the prior period (1990-1994), to expected number of claims assuming random claim arrival. In simulation of random claim arrival, paid claims are assumed to arrive at the *average* per-physician rate over 1995-2014; risk for each physician-year is: (1/20)*(4,694 actual claims over 1995-2014)/(15,881 physicians in balanced sample) = 0.01478. Actual counts over 1995-2014 are those observed in the data.

Paid claims during 1990-1994	No. of physicians	No. of physicians with Paid claims during the future (1995-2014) period								
		With 0 future paid claims			With 1+ future paid claims			With 2+ future paid claims		
		Actual	Expected	Actual/ Expected	Actual	Expected	Actual/ Expected	Actual	Expected	Actual/ Expected
0	14,440	11,768	10,735	1.10	2,672	3,705	0.72	662	494	1.34
1	1,174	700	886	0.79	474	288	1.65	174	33	5.27
2	206	86	141	0.61	120	65	1.85	62	5	12.4
3+	61	22	41	0.54	39	20	1.95	22	2	11
Total	**15,881**	**12,576**	**11803**	**1.07**	**3,305**	**4,078**	**0.81**	**920**	**534**	**1.72**

Consider first the columns for 0 future claims. Physicians with 0 prior-period claims are more likely to have 0 future-period claims than with random claim arrival. In contrast, physicians with one or more prior paid claims are less likely to have 0 future-period claims than if claims arrived at random. The larger the number of prior paid claims, the lower the actual/expected ratio. The columns for 1+ future claims show a consistent pattern: Physicians with 0 prior-period claims are *less* likely to have 1+ future-period claims than with random claim arrival. Conversely, physicians with one or more prior paid claims are *more* likely to have 0 future-period claims than if claims arrived at random.

Consider finally the columns for 2+ future claims. Many more physicians have 2+ future claims than would be expected if claims arrive at random. The actual/expected ratio increases monotonically with the number of prior-period claims. Vastly more physicians have 2+ future claims than expected with random claim arrival: 84 actual versus 7 expected. But why are there more actual than expected physicians with 0 prior-period claims, yet 2+ future claims? A plausible explanation is that the large pool of physicians with 0 paid claims includes many low-claim-risk physicians, but also some high-claim-risk physicians who, by good luck, had 0 prior-period paid claims. However, over the longer future period, their luck ran out, and they experienced multiple paid claims.

5. Single Failure Model

In this section we study the extent to which different physician characteristics predict paid claim risk. One important factor is paid claims during an initial five-year period, from 1990-1994 -- the first five years for which we have data. We study the 23,254 physicians who were licensed over 1990-1994, and divide them into groups based on the number of paid claims

they had in this period (0, 1, 2, or 3+ paid claims). The dataset includes only 61 physicians with 3+ claims during 1990-1994 (see Table 2), so it is not feasible to further divide this group into, say 3 versus 4+ prior claims. We then examine the paid claim risk of these physicians over the remaining 20 years in our sample period (1995-2014). In contrast to Table 2, we require physicians to be licensed only over 1990-1994, not over the entire sample period.

We also use this sample to study the extent to which other physician characteristics predict claim risk. The characteristics we study are: type of medical degree (M.D., D.O, or unknown); gender; practicing in a high-malpractice risk versus low-risk specialty; and medical school attended (U.S., foreign, or unknown).

A. Kaplan-Meier Survival Estimates

We next estimate the predictive power of paid claims during the prior period (1990-1994) for paid claims during the future period (1995-2014), using the nonparametric Kaplan-Meier survival model.

Figure 2 shows the Kaplan Meier survival probability over the future period for physicians with varying numbers of paid claims (0, 1, 2, and 3+) during the prior period. Shaded areas around each curve show 95% confidence intervals (C.I.'s). The differences in survival probability are visually apparent. For example, after 10 years, a physician with zero paid claims during 1990-1994 has an 88% chance of survival (no paid claims), compared to 68% for a physician with one prior paid claim, 49% for a physician with two prior paid claims, and 39% for a physician with 3+ prior paid claims.

There are similarly large differences in survival rates at longer horizons. However, at longer time horizons, the 95% C.I.'s for 2 versus 3+ prior claims overlap. The curves for physicians with prior paid claims are convex. This provides evidence that the predictive power of a prior claim wanes slowly over time.

Figure 2: Kaplan-Meier Survival Analysis for Physicians With 0, 1, 2, and 3+ Prior Paid Claims: Single-Failure Model

Kaplan Meier survival estimates, and 95% confidence intervals for fraction of Illinois physicians who "survive" (have no paid med mal claims) over a twenty-year future period (1995-2014). Figure shows separate curves physicians with 0, 1, 2, and 3+ paid claims during a five-year prior period (1990-1994). Sample is 23,254 Illinois physicians licensed throughout the prior period, of which 21,236 have 0 prior paid claims; 1,642 have 1 prior paid claim, 285 have 2 prior paid claims, and 91 have 3+ prior paid claims.

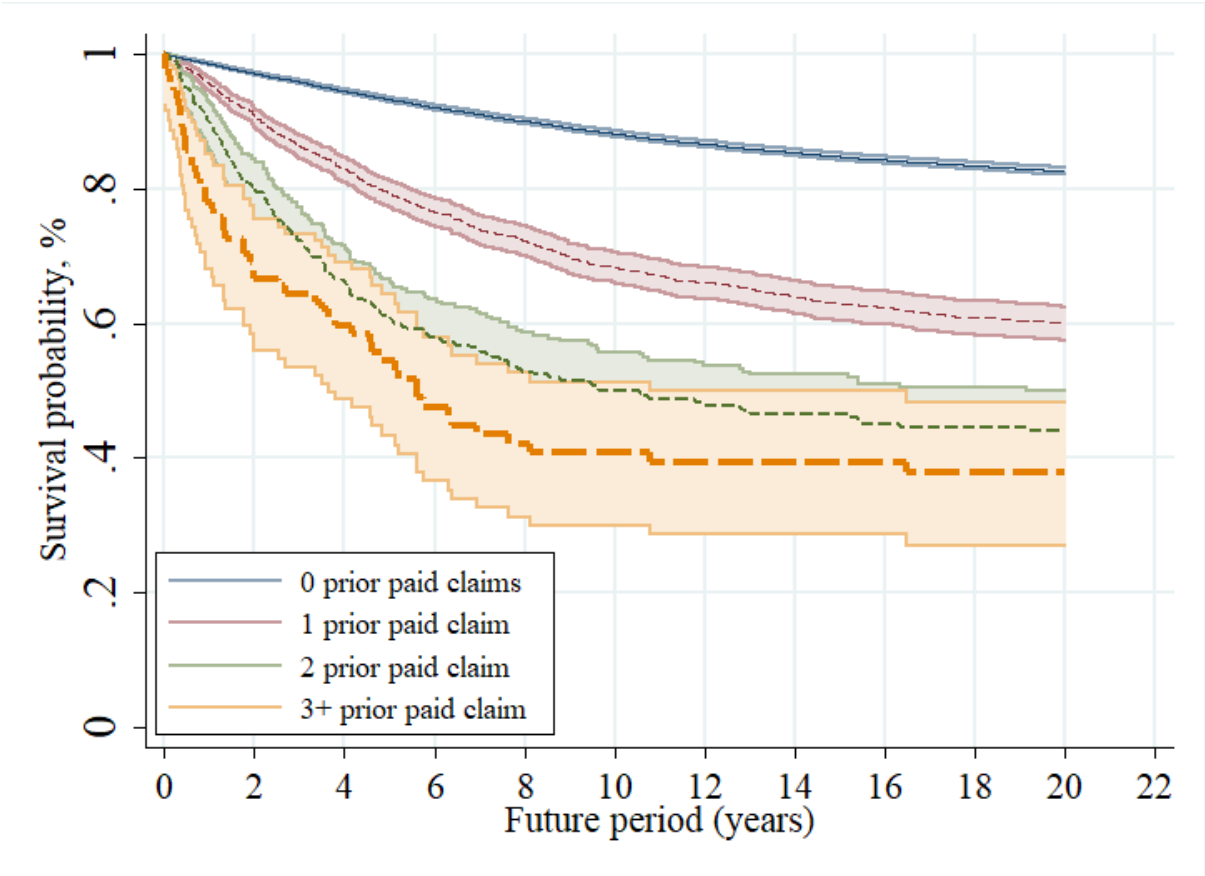


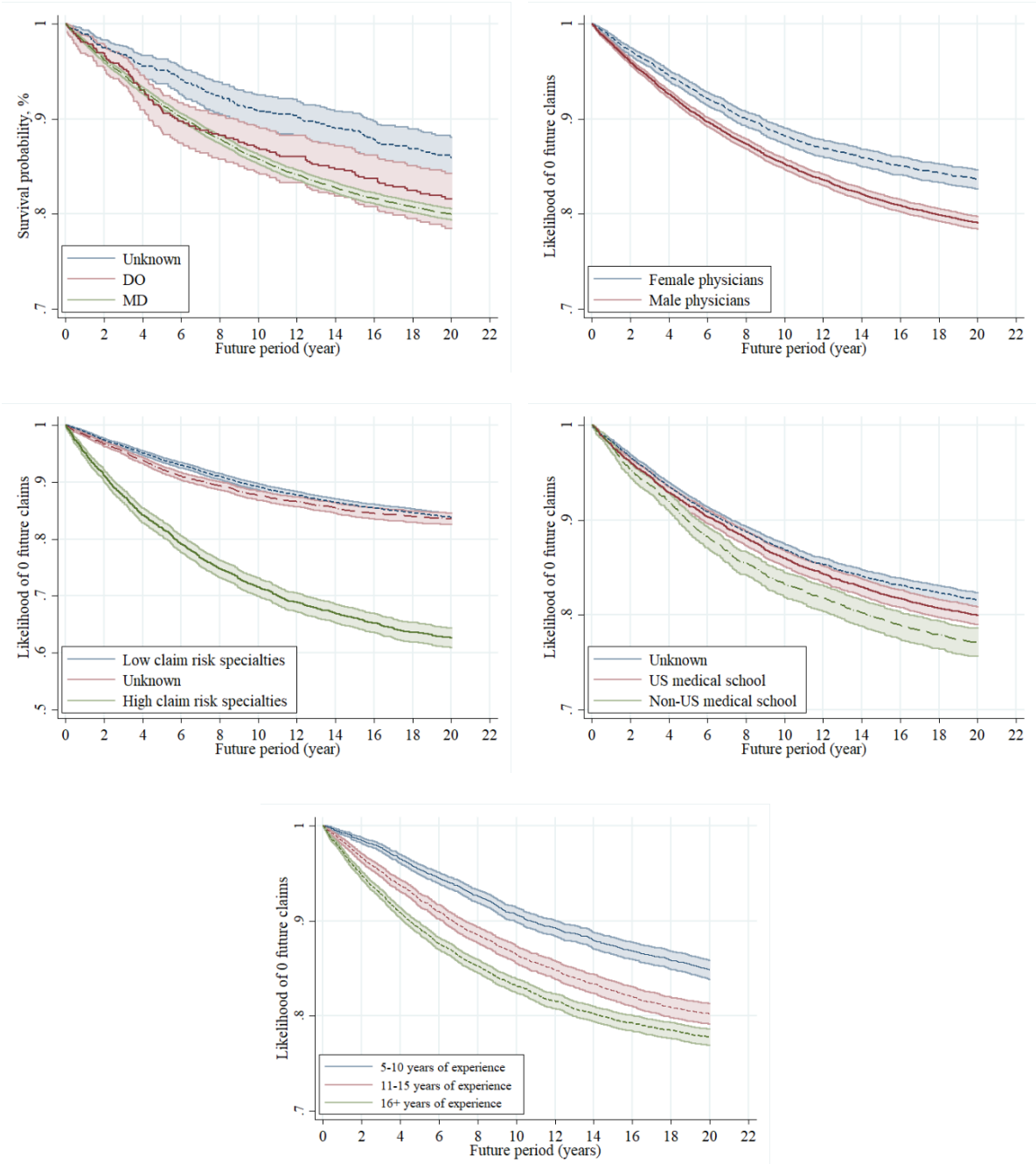
Figure 3 examines differences in survival rates by physician characteristics, both individually (Panel A) and interacted with the number of prior-period paid claims (Panel B).

Figure 3, Panel A: Survival Analysis by Physician Characteristics and Past Claim Record: Single-Failure Model

Kaplan-Meier survival estimates and 95% confidence intervals for survival (having zero paid claims) in the future period (1995-2014) by physician characteristics (Panel A) and both characteristics and prior paid claims (Panel B). Sample is same as Figure 2 (23,254 Illinois physicians licensed over 1990-1994).

Panel A. Survival Rates by Physician Characteristics

Top-left figure shows survival by degree type (D.O., M.D., or unknown). Top-right figure shows survival by gender (physicians with unknown gender are omitted); middle-left figure show survival based on specialty paid claim risk (low risk, high risk, unknown specialty); middle-right figure shows survival by location of medical school attended (U.S., Non-U.S., unknown); and bottom figure shows survival based on years of experience as of Dec. 31, 1994 (5-10 years, 11-15 years, and 16+ years). The vertical scale is different for different graphs.



Panel B. Survival Rates by Physician Characteristics and Prior Paid Claims

Survival rates (no paid claim) for the future period (1995-2014) by physician characteristics and paid claims in the prior period (1990-1994). Top-left figure shows survival by separately for M.D.'s and D.O.'s (physicians with unknown degree type are omitted) with versus without a prior period claim. Top-right figure shows survival by gender (physicians with unknown gender are omitted) and prior period claim; middle-left figure shows survival based on specialty (low risk or high risk, physicians with unknown specialty are omitted) and prior period claim; middle-right figure shows survival by medical school attended (U.S. or non-U.S., physicians with unknown medical school type are omitted) and prior period claim; and lower figure shows survival based on years of experience as of Dec. 31, 1994 (1-10 versus 11+ years) and prior period claim.

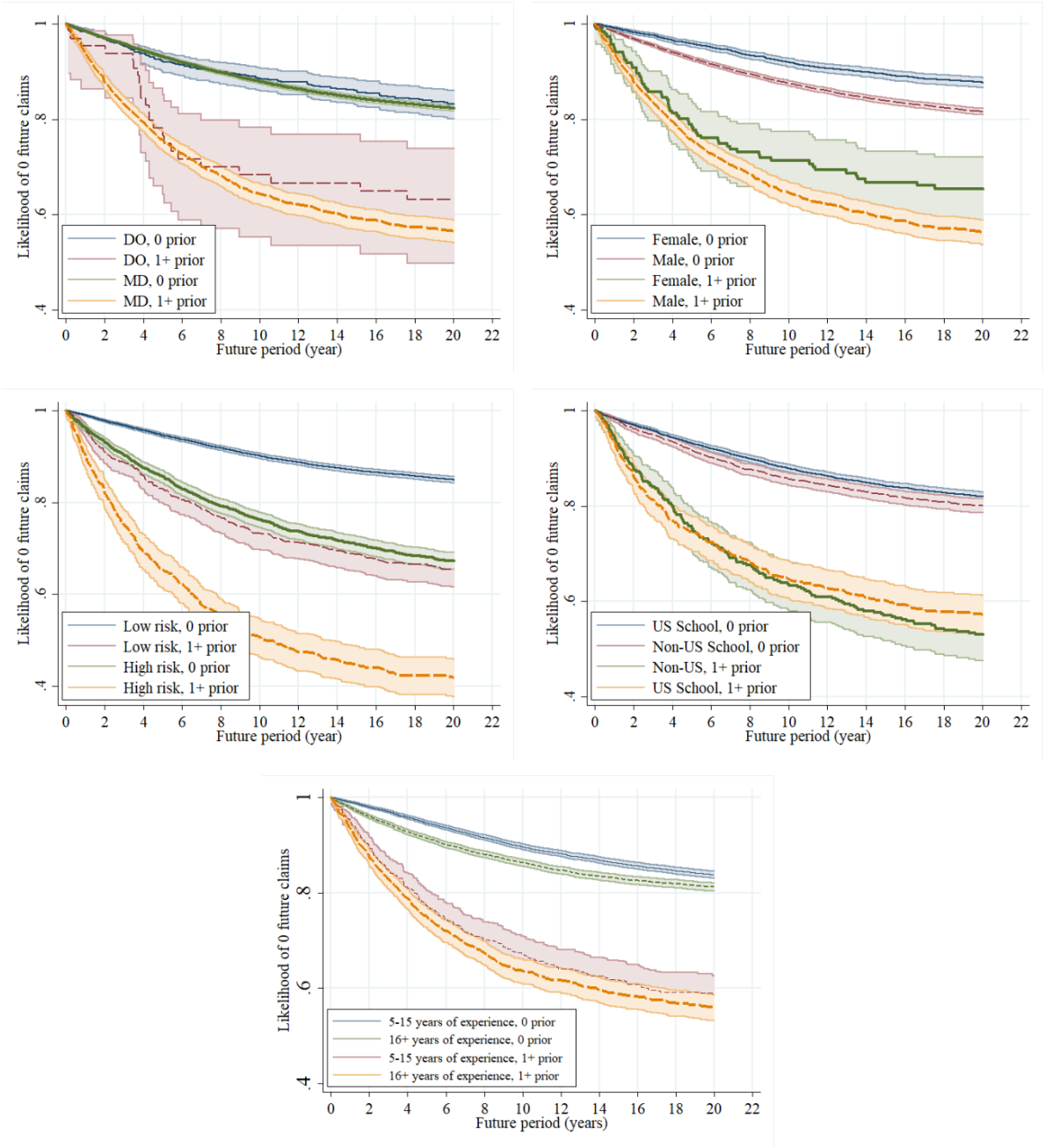


Figure 3 (Panel A) provides evidence on how survival rates without a paid claim vary with physician characteristics. The top-left graph shows survival rates by type of medical degree; physicians with unknown degree type are omitted. Survival rates for M.D. versus D.O. are similar for the first five years. There is mild evidence for M.D.'s having modestly higher

paid claim rates over longer periods. The confidence intervals overlap in Figure 3, but M.D.'s have significantly higher claim risk if we use the multiple failure model discussed in the next section. Physicians with unknown degree type have lower claim risk than either M.D.'s or D.O.'s. This provides evidence of selection bias in which physicians report their degree type.

The top-right graph shows survival rates separately for male and female physicians; physicians with unknown gender are omitted. Men face higher paid claim risk than women, but the differences are limited. For example, a female physician has about an 88% change of not having a paid claim over 10 years, versus 85% for a male physician. However, in the regression analysis below, the male-female difference increases if we control for specialty (men are more likely to practice in high-claim-risk specialties) and physician characteristics.

The middle-left graph shows survival rates separately for physicians in low-claim-risk specialties, high-claim-risk specialties, and those who did not report a specialty. High-claim-risk physicians face 3 times the claim risk as low-claim-risk physicians; for example, after 10 years, 11% of low-claim-risk physicians have experienced a claim, versus 29% of high-claim-risk physicians. Physicians without a reported specialty have claim risk similar to those in low-claim-risk specialties, but have lower claim risk in the multiple-failure analysis below. The greater paid-claim risk faced by physicians in high-claim-risk specialties is tautological, because we define high-versus-low paid claim risk based on our dataset, but the specialties we find to be high-claim-risk are similar to those found to be at high risk in other work (e.g., Jena et al., 2011; Studdert et al., 2016).

The middle-right graph shows survival rates separately for physicians who attended a US medical school, a non-US medical school, and those who did not report where they attended medical school. Physicians who attended a non-US school had somewhat higher claim risk than those who attended a US school (17% versus 14% over 10 years). Those who did not report school attended had slightly lower claim risk than those who attended a US medical school, suggesting a selection effect with regard to which physicians report the school they attended.

The bottom graph shows survival rates based on years of experience as of January 1, 1995 – the start of the future period. Future claim risk rises with prior experience.

In Figure 3, Panel B, we study the same physician characteristics as in Panel A, but interact each with indicators for whether a physician had either 0 paid claims in the prior period, or 1+ paid claims. The top-left figure shows survival rates for M.D. and D.O. physicians. Physicians with unknown degree type are omitted. For both physician types, those with 1+ prior claims have much higher claim rates in the future period than those without a prior period

claim. Within either the zero-prior-claim group or the 1+ prior claim group, M.D. and D.O. claim rates are similar.

The top-right graph shows survival rates by gender, separately for physicians with zero versus 1+ paid claims in the prior period. Physicians with unknown gender are omitted. Men face higher paid claim risk than women, but for both genders, physicians with 1+ prior paid claims have much higher risk of a paid claim than those without a prior paid claim. For example, a female physician with no prior claims has an 8% chance (92% survival) of having a paid claim over 10 years, while a female physician with 1+ prior paid claims has an 29% chance of having a paid claim over 10 years. For male physicians the corresponding percentages are 12% and 35%.

The middle-left figure shows survival rates by specialty type (high-claim -risk versus low-claim-risk specialty), again separately for physicians with zero versus 1+ prior paid claims. Within each group, physicians with 1+ prior period claims have much higher future claim risk. For example, low-claim-risk physicians with no prior claims have an 10% chance of having a paid claim over the next 10 years, versus 27% for low-claim-risk physicians with 1+ prior claims. High-claim-risk physicians with no prior claims have a 24% chance of having a paid claim over the next 10 years, versus 50% for high-claim-risk physicians with 1+ prior claims.

The middle-right figure shows survival rates by type of medical school attended (U.S. versus foreign), again separately for physicians with zero versus 1+ prior paid claims. Physicians whose medical school is unknown are omitted. Physicians with no prior paid claims who attended a US medical school have slightly lower paid claim rates than foreign-trained physicians with no prior claims. There are no significant differences in future claim rates based on medical school attended for physicians with 1+ prior paid claims. For both groups, there are large differences in survival rates for physicians with versus without a prior paid claim. For example, a U.S. trained physician with no prior paid claims has an 12% chance of having a paid claim over 10 years (88% survival), versus 25% for a U.S. trained physician with 1+ prior paid claims.

The bottom figure shows survival rates for less experienced (5-15 years versus more experienced (16+ years) physicians, with experience measured at the start of the future period. Both for physicians with and without a prior period claim, those with more experience have somewhat higher future claim rates, but the predictive effect of a prior claim is far greater than that for experience.

The common theme in the graphs in Panel B is the strong predictive power of a prior paid claim, within each of the subgroups. In contrast, the predictive power of M.D. versus D.O. and U.S. versus foreign medical school is small. There is somewhat more predictive power for gender and for years of experience, but still much less than the predictive power of prior paid claims.

B. Cox Model Estimates

We now turn to the semi-parametric Cox survival model, in which we study the predictive power of paid claims in the prior period, controlling for multiple physician characteristics (instead of one characteristic at a time, as in Figure 3, Panel B). We include in the model the characteristics that we studied in Figure 3.[6] We study physicians who were licensed over 1990-1994, but also require gender so that we can use gender as a covariate; the resulting sample is 21,526 physicians. Figure 4 presents survival curves for the base survival functions S_{0t} for physicians with different numbers of paid claims (0, 1, 2, and 3+) over 1990-1994.

[6] All physicians in the sample for this analysis are licensed during 1990-1994, so have at least 5 years of experience during the future period. We therefore use 5-10 years of experience as the omitted group for the Cox model estimates; the other included groups are 11-15 years and 16+ years of experience.

Figure 4. Hazard Analysis Using Cox Model: Single-Failure Model

Figure shows the base survival function S_{0t} for physicians to have paid claims over 1995-2014, based on the number of paid claims during a prior 5-year period (1990-1994). Sample is 21,526 physicians with imputed gender, licensed over 1990-1994. Covariates included in the Cox model are years of experience (using groups from Table 3) specialty (high risk, low-claim-risk, or unknown) medical school attended (US, foreign, or unknown), type of medical degree (M.D., D.O., or unknown), and gender.

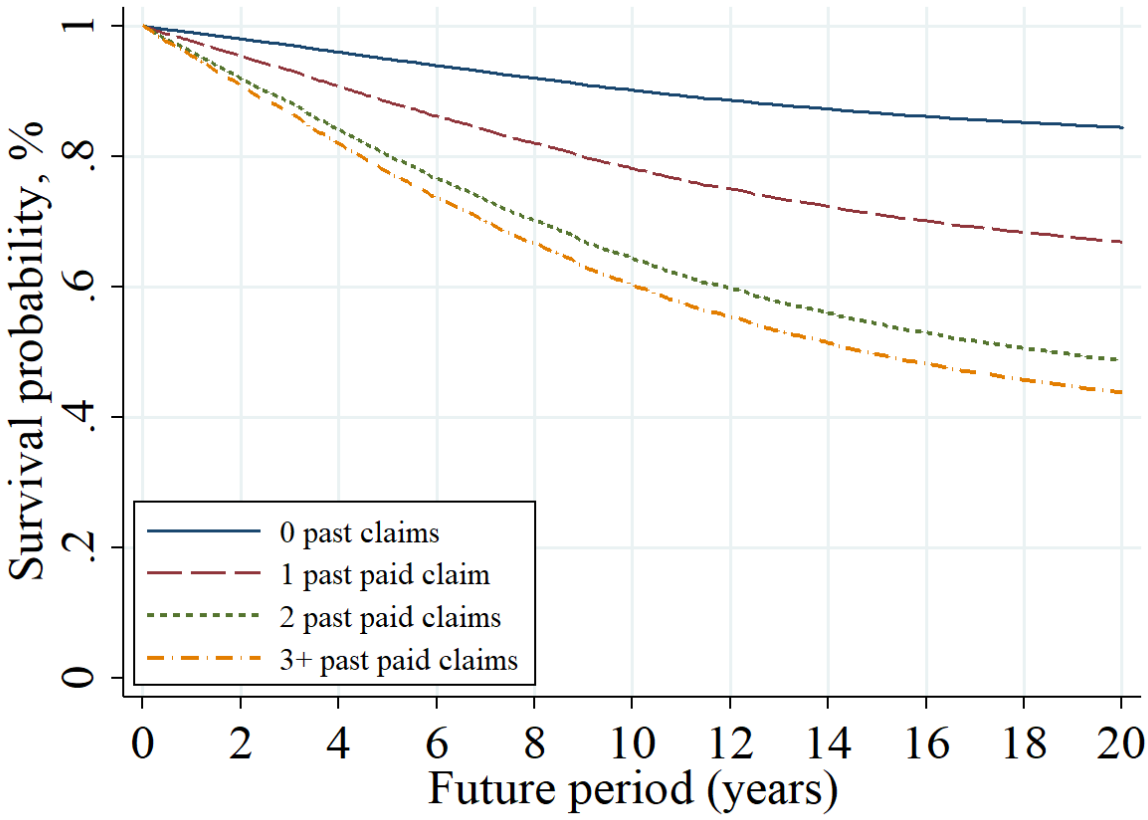


Figure 4 shows the strong tendency for physicians with more prior claims to have more future claims, after controlling for other physician characteristics. The survival rate after 10 years is 90% for physicians with no prior paid claims, versus 78% for those with one prior claim, 64% for those with two prior claims and 60% for those with 3+ prior claims.

C. Multivariate Regression Analysis: Single-Failure Model

We turn next to regression analysis, using the Cox model and the same covariates as in Figure 4. In the left-hand columns of Table 3, the sample is physicians licensed during the prior period (1990-1994). In columns (1)-(3) we include the number of prior period claims as a predictor. In columns (4)-(6), we instead use dummy variables for 1 prior claim, 2 prior claims, and 3+ prior claims (no prior claims is the omitted group). We report hazard ratios for the covariates that we studied individual in Figure 3. The omitted groups are low-claim-risk specialty, D.O. degree, U.S. medical school and male gender (for regressions that include

gender). The right-hand columns are similar, except that the sample is physicians licensed over 1990-2014.

Table 3. Factors Predicting Paid Claim Hazard Rates, Single-Failure Model

Note: Single-failure model. Sample is physicians licensed over 1990-1994 (columns 1-6) or physicians licensed over 1990-2014 (columns 7-12). Prior claims are counted during the initial period (1990-1994). Dependent variable is survival time. Omitted categories are: low-claim-risk specialty; DO degree, low-claim-risk specialty, attended U.S. medical school, and 6-10 years of experience (experience is measured as of Jan. 1, 1995).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
No. of prior claims	1.88***	1.82***	1.88***				1.77***	1.71***	1.79***			
	(0.06)	(0.06)	(0.06)				(0.06)	(0.06)	(0.07)			
1 prior claim				2.29***	2.18***	2.27***				2.09***	1.99***	2.11***
				(0.11)	(0.10)	(0.11)				(0.11)	(0.10)	(0.12)
2 prior claims				3.46***	3.25***	3.59***				3.14***	2.93***	3.25***
				(0.34)	(0.32)	(0.36)				(0.34)	(0.32)	(0.37)
3+ prior claims				4.47***	4.21***	4.38***				3.64***	3.40***	3.81***
				(0.78)	(0.73)	(0.83)				(0.72)	(0.66)	(0.81)
High claim-risk specialty	2.31***		2.25***	2.30***		2.24***	2.39***		2.31***	2.38***		2.30***
	(0.09)		(0.09)	(0.09)		(0.09)	(0.10)		(0.10)	(0.10)		(0.10)
Unknown specialty	0.99		1.02	0.99		1.02	1.11**		1.15**	1.11**		1.15**
	(0.04)		(0.04)	(0.04)		(0.04)	(0.06)		(0.06)	(0.06)		(0.06)
specialty dummies		Yes			Yes			Yes			Yes	
MD	0.94	1.05	0.97	0.94	1.05	0.97	0.93	1.05	0.97	0.94	1.06	0.97
	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)	(0.09)	(0.10)	(0.11)	(0.10)	(0.09)	(0.11)	(0.10)
Unknown degree	0.73**	0.81	0.78*	0.73**	0.82	0.78*	0.73**	0.83	0.78*	0.74**	0.83	0.79
	(0.09)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.12)	(0.11)	(0.10)	(0.12)	(0.11)
Non-U.S. medical school	1.17***	1.18***	1.18***	1.18***	1.19***	1.19***	1.17***	1.18***	1.17***	1.18***	1.19***	1.17***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Unknown medical school	0.96	0.96	0.94	0.96	0.96	0.94	0.97	0.98	0.96	0.98	0.98	0.97
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Female			0.72***			0.72***			0.71***			0.71***
			(0.03)			(0.03)			(0.04)			(0.04)
11-15 years experience	1.28***	1.28***	1.25***	1.26***	1.26***	1.23***	1.23***	1.23***	1.20***	1.21***	1.22***	1.18***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
16+ years experience	1.38***	1.43***	1.27***	1.35***	1.40***	1.25***	1.35***	1.40***	1.23***	1.33***	1.38***	1.22***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)
No. of physicians	23,254	23,254	21,526	23,254	23,254	21,526	15,881	15,881	14,710	15,881	15,881	14,710
No. of claims	3,957	3,957	3,585	3,957	3,957	3,585	3,305	3,305	3,006	3,305	3,305	3,006

The regressions in columns (1)-(3) indicate that on average, each additional prior med mal claim roughly doubles the risk of a future claim. The decomposition based on number of paid claims in columns (4)-(6) shows that the largest jump in future claim risk is for physicians with one versus zero prior paid claims, with a hazard ratio is around 2.4. Physicians with two prior claims have about 3.7 times the future claim risk as those with no prior claims, and about

1.5 times the future claim risk as those with one prior claim.[7] Physicians with 3+ prior claims have about 4.5-4.8 times the future claim risk as those with no prior claims, and about 1.3 times higher risk than those with two prior claims.[8]

We turn next to the hazard ratios for covariates, and consider first column (1). High-risk-claim specialties have roughly 2.3 times higher claim risk than low-risk specialties. Physicians without a reported specialty have risk similar to those in known, low-risk specialties. Hazard rates are similar for M.D.'s and D.O.'s. The lower risk for physicians with unknown degree likely reflects self-selection for those physicians who do not report degree type. Foreign-trained physicians have roughly 1.2 times higher claim risk than U.S. trained physicians. However, we find in separate work (Hyman, Rahmati and Black, 2021) that foreign-trained physicians are not more likely to face medical board discipline than U.S. trained physicians. Thus, their higher paid claim risk could reflect patients being more willing to sue foreign-trained physicians rather than these physicians being more likely to provide substandard care.

For experience, all physicians considered in Table 3 have at least 5 years of experience as of January 1, 1995 – the start of the future period. Physicians with 11-15 years of experience have higher paid claim rates than the omitted group (6-10 years) experience, and physicians with 16+ years of experience have slightly higher paid claim rates than those with 11-15 years of experience.

In column (2), we remove the high-claim-risk and low-claim-risk categories and include specialty dummies. The coefficients for type of medical degree move somewhat, and unknown degree becomes only marginally significant. The coefficients for medical school attended are little changed. In column (3), we return to using the claim risk groups, drop physicians with unknown gender, and report the coefficient on female gender. Female physicians have around 70% of the male hazard rate, controlling for other predictors. Finally, in columns (4)-(6) we switch from number of prior claims as a single variable to dummy variables for different numbers of prior claims. The coefficients on covariates are very similar to those in columns (1)-(3).

[7] For example, the ratio of hazard ratios for two prior paid claims versus one prior paid claim in column (1) is 3.74/2.44 = 1.53.

[8] The ratio of hazard ratios for 3+ versus two prior paid claims in column (1) is 4.80/3.74 = 1.28.

We turn in columns (7)-(12) to the balanced sample of physicians who were licensed in each year over 1990-2014. The predictive power of prior claims is modestly reduced but remains strong. The coefficients on covariates are very similar to those in columns (1)-(6).

6. Multiple Failure Model

In Part 5, we relied on a single-failure survival model, in which: (i) physicians are classified based on number of claims during the first five years of the sample period; and (ii) during the future period we study only the first paid claim, which we treat as non-survival. We switch here to a dynamic survival model (Anderson-Gill, 1982) which allows us to study multiple claims against the same physician. This approach does not require us to specify an arbitrary initial period during which we count prior claims or to limit the analysis to physicians licensed during that prior period. At each time that a claim occurs for a particular physician, we count how many prior claims this physician had, together with the physician's experience years at the time of the claim. In analyses which control for the number of prior claims, all physicians initially begin in a zero-claim group. When a physician experiences a first paid claim, we treat this as nonsurvival within the zero-claim group, move the physician to a one-prior-claim group, and reset the time counter to zero, thus treating the physician as newly entering the data with one prior paid claim. When a physician in the one-prior-claim group experiences a second paid claim, we treat this as nonsurvival within the one-prior-claim group, move the physician to a two-prior-claim group, and again reset the time counter to zero, thus treating the physician as newly entering the data with two prior paid claims. Physicians leave the sample when they are no longer licensed.

An important limitation of the multiple-failure model is that, within each group (based on number of prior claims), it treats all paid claims (failures) as independent. Consider, for example, the one-prior-claim group. The model cannot distinguish between physicians who had a first prior claim early on (relative to when the physician enters the sample), versus those who had a much longer period with no claims.

A. Univariate Estimates

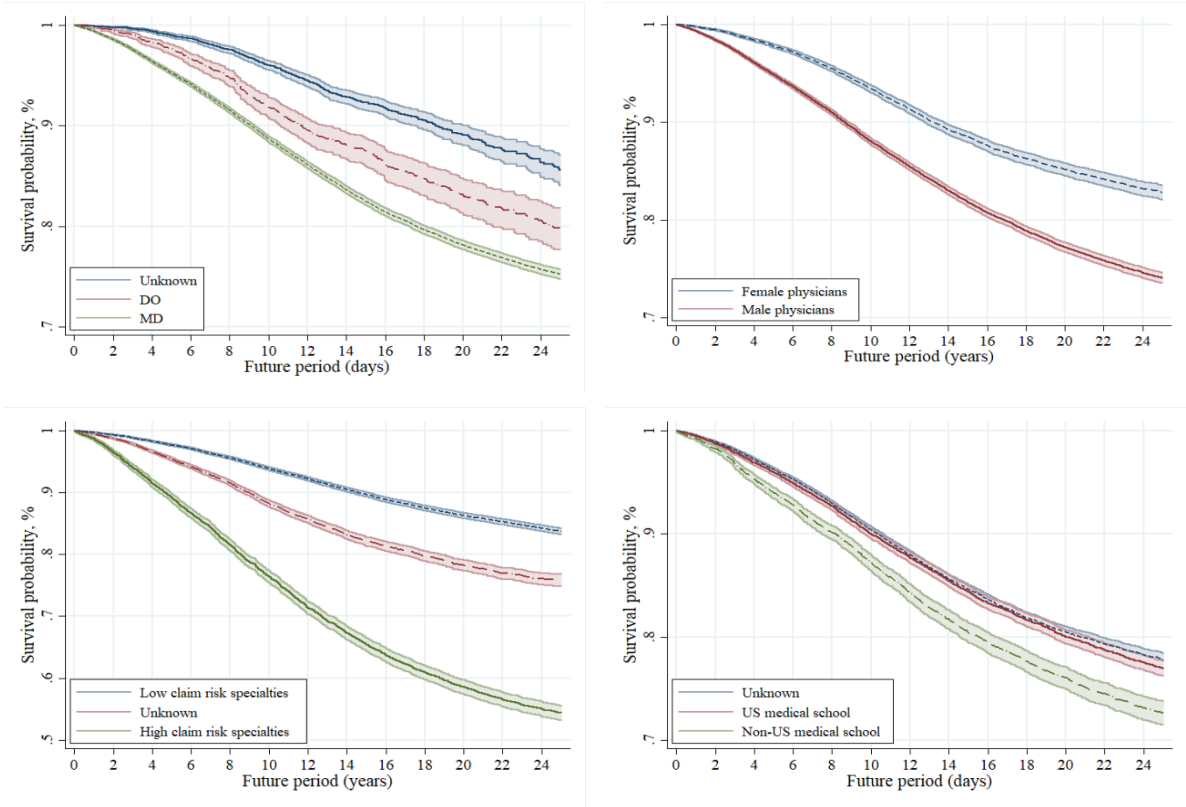
We begin with multiple-failure graphs, which are otherwise similar to the single-failure graphs in Figure 3. Figure 5 shows how survival rates vary with physician characteristics: type of medical degree, gender, specialty, location of medical school attended, and experience.

Figure 5: Kaplan Meier Survival Rates by Physician Characteristics and Past Claims: Multiple-Failure Model

Kaplan-Meier survival estimates and 95% confidence intervals for survival (having zero paid claims) by physician characteristics (Panel A) and both characteristics and prior paid claims (Panel B), using multiple failure model. We compute the number of prior paid claims from 1990 or licensure (whichever is later) through the day prior to the claim of interest. Sample is 84,114 physicians, licensed in Illinois at any time over 1990-2014.

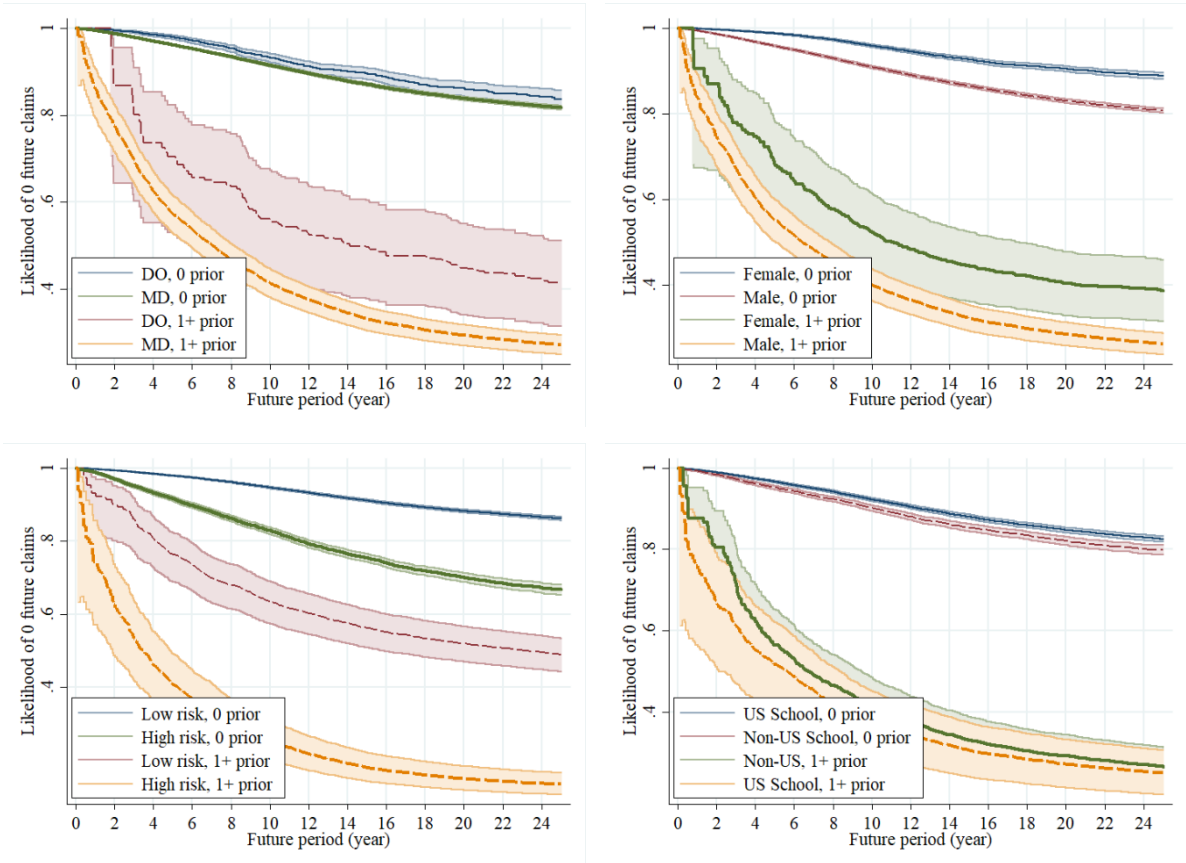
Panel A. Survival Rates by Physician Characteristics

Top-left figure shows survival by degree type (D.O., M.D., or unknown). Top-right figure shows survival by gender (physicians with unknown gender are omitted); bottom-left figure show survival based on specialty paid claim risk (low risk, high risk, unknown specialty); and bottom-right figure shows survival by location of medical school attended (U.S., Non-U.S., unknown). Y-axis scale differs across graphs.



Panel B. Survival Rates by Physician Characteristics and Prior Paid Claims

Survival rates (no paid claim) by physician characteristics and prior claim history. Top-left figure shows survival by existence of a prior paid claim and degree type (D.O. or M.D., unknown degree type is omitted). Top-right figure shows survival by existence of prior claim and gender (unknown gender is omitted); bottom-left figure shows survival based on existence of prior claim and specialty (low risk or high risk, unknown specialty is omitted); and bottom-right figure shows survival by existence of prior claim and medical school attended (U.S. or non-U.S., unknown is omitted).



Comparing the multiple-failure graphs in Figure 5, Panel A to the comparable single-failure graphs in Figure 3, Panel A, we obtain qualitatively similar results but tighter confidence intervals and sometimes clearer separation between groups. For type of medical degree, MDs again show the highest claim risk, then DOs, then physicians with unknown degree type. Male physicians have higher claim risk than female physicians. As expected, high-claim-risk specialties have higher claim risk than low-claim-risk specialties, but physicians with unknown specialty clearly face intermediate claim risk, in contrast to the single-failure graphs where physicians in low-risk specialties and those with unknown specialty faced similar claim risk. For type of medical school attended, physicians who attended U.S. medical schools face claim risk very similar to those whose school type is unknown; while physicians who attended non-U.S. schools face higher claim risk.

Comparing the multiple-failure graphs separated by prior claim history in Figure 5, Panel B to the comparable single-failure graphs in Figure 3, Panel B, we again obtain

qualitatively similar results but tighter confidence intervals and sometimes clearer separation between groups. For example, in the multiple failure model, MDs with 1+ prior claims have higher claim risk than DOs with 1+ prior claims; in the single-failure model, in contrast, the confidence intervals overlap.

B. Multivariate Regression Estimates: Multiple-Failure Model

We turn in Table 4 to multivariate regression estimates using the multiple-failure model. The regression models in Table 4 are the same as in Table 3, but the sample is much larger because it is not limited to physicians licensed throughout 1990-1994. Experience is measured at the time of each paid claim**.** Table 4, regressions (1)- (3) indicate that on average, each additional prior paid claim increases the hazard rate for a future paid claim by roughly 50%.[9] In regressions (4)-(6), we switch to dummies for 1, 2 and 3+ prior paid claims; the omitted group is physicians with zero prior paid claims. These regressions indicate that an increase from zero to one prior paid claim roughly quadruples the hazard rate for a future paid claim. Physicians with two prior paid claims have roughly six times higher risk of a future paid claim as physicians with no prior claims, and roughly 50% higher risk than physicians with one prior claim. Physicians with three or more prior paid claims have roughly eight times the risk of a future paid claim, relative to physicians with no prior claims, and roughly one-third higher risk than those with two prior claims, and twice the risk of those with one prior claim. The overall pattern is that each additional prior paid claim predicts higher future claim risk, but at a decreasing rate – the most powerful relative predictor is having one prior paid claim versus none. In Tables A-5 and A-6, we report the results of similar analyses using Gompertz and Weibull models, and confirm that results are similar across all three models.

[9] If we winsorize the number of prior paid claims at 3, which is more closely comparable to the dummy variable approach in cols. (4)-(6), the coefficients in cols. (1)-(3) rise to approximately 2.2.

Table 4. Factors Predicting Paid Claim Hazard Rates, Multiple Failure Model

Note: Table uses multiple failure model, and shows hazard rate of future paid claims as a function of prior paid claims plus other covariates, using different hazard models and various dummies. We compute the number of prior paid claims from 1990 or licensure (whichever is later) prior to the claim of interest. In regressions (1)-(3), we use the number of paid claims, while in columns (4)-(6), we use dummies for 1, 2, and 3+ past claims. Experience is determined from license issuance to the date of claim. Standard errors in parentheses. Omitted categories for columns (1) and (4) are: 0-5 years of experience; low-claim-risk specialty; DO degree, low-claim-risk specialty, and attended U.S. medical school. In columns (4-(6), 0 prior claims is the omitted category. In columns (2) and (5) we include specialty dummies, the omitted specialty is family medicine. Columns (3) and (6) omit persons with unknown gender, male is the omitted gender.

	(1)	(2)	(3)	(4)	(5)	(6)
Number of prior claims	1.54***	1.53***	1.53***			
	(0.03)	(0.03)	(0.03)			
1 prior claim				4.14***	3.98***	4.02***
				(0.13)	(0.13)	(0.13)
2 prior claims				6.00***	5.74***	5.79***
				(0.29)	(0.28)	(0.29)
3+ prior claims				8.32***	7.76***	8.15***
				(0.59)	(0.56)	(0.61)
High claim-risk specialty	3.06***		2.91***	2.63***		2.52***
	(0.10)		(0.11)	(0.07)		(0.07)
specialty dummies		Yes			Yes	
Unknown specialty	1.71***	1.29***	1.72***	1.62***	1.31***	1.62***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.05)
M.D.	1.25***	1.36***	1.25***	1.28***	1.37***	1.27***
	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)
Unknown medical degree	0.51***	0.55***	0.54***	0.55***	0.59***	0.58***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Non-U.S. medical school	1.34***	1.37***	1.30***	1.27***	1.30***	1.24***
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
Unknown medical school	0.97	0.97***	0.96	0.95*	0.96***	0.95**
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
Female			0.52***			0.54***
			(0.02)			(0.02)
6-10 years of experience	2.96***	2.94**	2.87***	3.04***	3.03***	2.94***
	(0.24)	(0.24)	(0.25)	(0.25)	(0.25)	(0.25)
11-15 years of experience	3.17***	3.17***	2.91***	3.23***	3.21***	2.96***
	(0.27)	(2.69)	(0.26)	(0.27)	(0.27)	(0.26)
16+ years of experience	1.80***	1.84***	1.50***	1.61***	1.63***	1.35***
	(0.15)	(0.16)	(0.13)	(0.14)	(0.14)	(0.12)
No. of physicians	84,114	84,114	76,925	84,114	84,114	76,925
No. of claims	9,827	9,827	8,932	9,827	9,827	8,932

In Figure 6, we juxtapose survival estimates for two types of multiple failure models: a univariate Kaplan-Meier model (left-side figure) and a multi-variate Cox-model (right-side figure). Compared to the comparable single failure estimates in Figure 2 and Figure 4, the period covered is longer (25 years instead of 20 years) because with the multiple-failure model we can use the full sample period to estimate survival probabilities. Both models in Figure 6

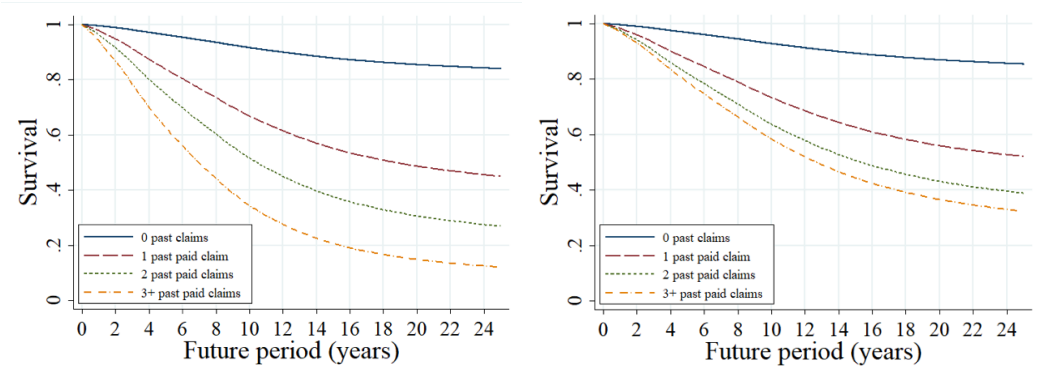
show survival probabilities over the sample period, for physicians with varying numbers of prior paid claims (0, 1, 2, and 3+).[10]

As was the case for the single-failure model, the differences in survival probability are visually apparent. With the Kaplan-Meier model (left-side figure), after 10 years, a physician with zero prior paid claims has a 92% chance of survival (no paid claims), compared to 66% for a physician with one prior paid claim, 52% for a physician with two prior paid claims, and 34% for a physician with 3+ prior paid claims. The differences in survival rates generally increase over time. The survival curves are convex. This provides evidence that the predictive power of a prior claim wanes over time, although slowly.

The Cox model in the right-side figure adds controls for the covariates in Table 4, column 4. Although the overall pattern is similar to our findings with the Kaplan-Meier model (i.e., a higher number of prior claims predicts a lower future survival curve), we find a higher likelihood of survival when we control for covariates. For example, with the Cox model (right-side figure), after 10 years, a physician with zero prior paid claims has a 93% chance of survival (no paid claims), compared to 73% for a physician with one prior paid claim, 63% for a physician with two prior paid claims, and 58% for a physician with 3+ prior paid claims.

Figure 6: Kaplan-Meier (left-side) and Cox Model Survival Analysis for Physicians with 0, 1, 2, and 3+ Past Claims: Multiple-Failure Model

Kaplan Meier (left-side) and Cox model (right-side) survival estimates for fraction of Illinois physicians from 1990-2014. Covariates for the Cox model (other than number of prior claims) are same as in Table 4, col. (4). Figure shows separate curves for survival as separated by number of paid claims, with values 0, 1, 2, and 3+ paid claims. Sample is 84,114 Illinois physicians licensed at any time during 1990-2014.



C. Estimates Limited to Physicians with Paid Claims

Many prior studies of malpractice claiming, including the related Studdert et al. (2016) analysis have relied on the NPDB for data. The NPDB, however, contains data only on

[10] Figure 6 does not include confidence intervals, since we were not able to find a way to plot them in Stata with a multiple-failure model.

physicians with paid claims. We examine in this section how imposing a similar restriction on our data would affect the results reported above. Table 5 presents results for the multiple failure model, for Illinois physicians with at least one paid claim during our sample period. For columns (4)-(6), the omitted reference group is physicians with one prior claim.

When restricting to physicians with at least one prior claim, one loses important information about the predictive value of a first paid claim. Having two paid claims predicts higher future risk than having one prior claim, although the relative increase is higher for the restricted sample in Table 5 (for example, the coefficient for 2 prior claims in Table 5 is around 3, versus a ratio for two versus one prior claims in Table 4 of around 1.5). A second major difference between Tables 4 and 5 lies in the coefficients on the experience dummies. In table 4, claim risk is highest for 6-15 years of experience. In Table 5, in contrast, risk of an additional claim falls with additional experience years. This could reflect two effects: First, the longer until one's first claim, the lower the risk for a second claim. We test this directly by studying whether for physicians with one or more claims, the time from entry into the sample until the first claim predicts the likelihood of a second claim. We confirm, for physicians who remain in the sample for at least 5 years after a first claim, time to first claim strongly predicts lower likelihood of a second claim within this 5-year period.[11] Second, physicians may change their behavior in response to having a first paid claim. We cannot test this second explanation.

[11] In a regression of a dummy variable for existence of a second claim on time to first claim and our other covariates (except experience years), an extra year before the first claim predicts 0.3% lower likelihood of a second claim ($t = 34.08$).

Table 5. Results with Multiple Claims Per Physician with at Least One Claim, Cox Model

Note: Table uses multiple failure model, and is similar to Table 4, except sample is limited to physicians with at least one paid claim during our sample period.

	(1)	(2)	(3)	(4)	(5)	(6)
Number of prior claims	1.41***	1.40***	1.42***			
	(0.02)	(0.02)	(0.02)			
2 prior claims				3.10***	3.07***	3.08***
				(0.16)	(0.16)	(0.17)
3+ prior claims				4.42***	4.25***	4.44***
				(0.34)	(0.32)	(0.36)
High claim-risk specialty	2.48***		2.44***	2.21***		2.17***
	(0.12)		(0.13)	(0.11)		(0.12)
specialty dummies		Yes			Yes	
Unknown specialty	2.20***	2.05***	2.22***	2.06***	2.00***	2.06***
	(0.12)	(0.21)	(0.13)	(0.11)	(0.18)	(0.12)
M.D.	1.41***	1.45***	1.44***	1.42***	1.43***	1.44***
	(0.17)	(0.17)	(0.17)	(0.16)	(0.16)	(0.16)
Unknown medical degree	0.77	0.80	0.84	0.82	0.85	0.90
	(0.13)	(0.13)	(0.14)	(0.13)	(0.14)	(0.15)
Non-U.S. medical school	1.23***	1.27***	1.22***	1.18***	1.22***	1.17**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Unknown medical school	1.08*	1.09*	1.06	1.05	1.06	1.03
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Female			0.62***			0.65***
			(0.04)			(0.04)
6-10 years of experience	0.82	0.84	0.72	0.81	0.82	0.70
	(0.32)	(0.34)	(0.28)	(0.30)	(0.32)	(0.26)
11-15 years of experience	0.57	0.58	0.49*	0.57	0.57	0.49*
	(0.23)	(0.24)	(0.20)	(0.22)	(0.23)	(0.19)
16+ years of experience	0.21***	0.22***	0.17***	0.19***	0.20***	0.16***
	(0.08)	(0.09)	(0.07)	(0.08)	(0.08)	(0.06)
No. of physicians	6,857	6,268	6,857	6,857	6,268	6,857
No. of claims	3,270	2,941	3,270	3,270	2,941	3,270

7. Discussion

Using more detailed and granular data than has previously been available, we study the impact of physician characteristics on the hazard rate of paid med mal claims. Our data is limited to one state (Illinois), but includes all licensed physicians in that state, including those with zero paid claims. Prior studies have relied on the NPDB, which only has data on physicians with 1+ paid claims. We also have information on gender, specialty, years of experience, type of degree (M.D. versus D.O.), country of medical school attendance (U.S. versus non-U.S.).

After controlling for other factors, we find that physicians with a single prior paid claim have a four-fold higher risk of future claims compared to physicians with zero prior paid claims – and each incremental prior paid claim is associated with higher future claim risk. Male gender, attending a non-U.S. medical school, having an M.D. rather than D.O. degree (in multiple-failure but not single failure models), and practicing in a high-malpractice-risk specialty all predict higher paid claim risk. Paid claim risk is also higher for physicians with 6-15 prior

years of experience than for those who are either earlier or later in their careers. We obtain comparable results whether we use single-failure or multiple-failure models, and whether we use Kaplan-Meier, Cox, Gompertz, or Weibull models.

8. Conclusion

Physicians often claim that med mal claims are random events, like being struck by lightning. We find that med mal claims are not random. Certain physician characteristics are consistently associated with higher (or lower) med mal claim risk. These results are robust to multiple alternative models. Although we have only imperfect measures of skill (i.e., experience and board certification), our findings make it clear that med mal claiming is not random.

References

Adamson, T. E., Baldwin Jr, D. C., Sheehan, T. J., & Oppenberg, A. A. (1997). Characteristics of surgeons with high and low malpractice claims rates. Western Journal of Medicine, 166(1), 37.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. The annals of statistics, 1100-1120.

Black, B., Hyman, D. A., & Lerner, J. Y. (2019). Physicians with multiple paid medical malpractice claims: Are they outliers or just unlucky?. *International Review of Law and Economics*, *58*, 146-157.

Bovbjerg, R. R., & Petronis, K. R. (1994). The relationship between physicians' malpractice claims history and later claims: does the past predict the future?. *Journal of the American Medical Association*, *272*(18), 1421-1426.

Daughety, A. F., & Reinganum, J. F. (2017). *Settlement and trial* (Vol. 3, pp. 229-246). Oxford: Oxford University Press.

Ely, J. W., Dawson, J. D., Young, P. R., Doebbeling, B. N., Goerdt, C. J., Elder, N. C., & Olick, R. S. (1999). Malpractice claims against family physicians: are the best doctors sued more?. *Obstetrical & gynecological survey*, *54*(9), 561-562.

Hyman, D.A., Rahmati, M., and Black, B. (2021), Medical Malpractice and Physician Discipline: The Good, The Bad and The Ugly, *Journal of Empirical Legal Studies* 18, 131-166.

Jena, Anupam B., Seabury, Seth, Lakdawalla, Darius, & Chandra, Amitabh. (2011). Malpractice risk according to physician specialty. *New England Journal of Medicine*, *365*(7), 629-636.

Paik, Myungho, Bernard Black, and David Hyman (2013), The Receding Tide of Medical Malpractice Litigation Part 1: National Trends, *Journal of Empirical Legal Studies* 10, 612-638

Pukk-Härenstam, K., Ask, J., Brommels, M., Thor, J., Penaloza, R. V., & Gaffney, F. A. (2008). Analysis of 23, 364 patient-generated, physician-reviewed malpractice claims from a non-tort, blame-free, national patient insurance system: lessons learned from Sweden. *BMJ Quality & Safety*, *17*(4), 259-263.

Sloan, F. A., Mergenhagen, P. M., Burfield, W. B., Bovbjerg, R. R., & Hassan, M. (1989). Medical malpractice experience of physicians: predictable or haphazard?. *Jama*, *262*(23), 3291-3297.

Studdert, D. M., Bismark, M. M., Mello, M. M., Singh, H., & Spittal, M. J. (2016). Prevalence and characteristics of physicians prone to malpr2016)actice claims. *New England Journal of Medicine*, *374*(4), 354-362.

Taragin, M. I., Wilczek, A. P., Karns, M. E., Trout, R., & Carson, J. L. (1992). Physician demographics and the risk of medical malpractice. *The American journal of medicine*, *93*(5), 537-542.

Unwin, E., Woolf, K., Wadlow, C., Potts, H. W., & Dacre, J. (2015). Sex differences in medico-legal action against doctors: a systematic review and meta-analysis. *BMC Medicine*, *13*(1), 172.

Venezian, E. C., Nye, B. F., & Hofflander, A. E. (1989). The distribution of claims for professional malpractice: some statistical and public policy aspects. *Journal of Risk and Insurance*, 686-701.

Waters, T. M., Lefevre, F. V., & Budetti, P. P. (2003). Medical school attended as a predictor of medical malpractice claims. *BMJ Quality & Safety*, *12*(5), 330-336.

Watanabe, Y. (2006). Learning and bargaining in dispute resolution: Theory and evidence from medical malpractice litigation. *Unpublished manuscript*.

Yousefi, K. (2018). The More Med-Mals, the Shorter the Litigation: Evidence from Florida. *Review of Law & Economics*, *14*(1).