Health Economics, Policy and Law (2024), page 1 of 15
doi:10.1017/S174413312400015X

HEALTH ECONOMICS, POLICY and LAW

ARTICLE

The inefficient effects of non-clinical factors on health care costs

Shawn McFarland and Jonathan Miller

Department of Finance, Insurance and Real Estate, University of Memphis, Memphis, TN, USA
Corresponding author: Shawn McFarland; Email: smmcfrl1@memphis.edu

(Received 31 July 2023; revised 14 June 2024; accepted 16 August 2024)

Abstract
We use Benford's law to examine the non-random elements of health care costs. We find that as health care expenditures increase, the conformity to the expected distribution of naturally occurring numbers worsens, indicating a tendency towards inefficient treatment. Government insurers follow Benford's law better than private insurers indicating more efficient treatment. Surprisingly, self-insured patients suffer the most from non-clinical cost factors. We suggest that cost saving efforts to reduce non-clinical expenses should be focused on more severe, costly encounters. Doing so focuses cost reduction efforts on less than 10% of encounters that constitute over 70% of dollars spent on health care treatment.

Keywords: cost reduction; health care; insurance; policy

JEL Codes: I10; I13; I18; G22; G52

Many citizens are unhappy about rising health care costs in the United States, and while patients' diseases and illnesses are random, their cost of treatment is not. The variety, severity, and duration of symptoms varies across individuals, based largely on the individual's prior health condition, underlying comorbidities, and genetic makeup. The individual nature of each health care event requires individualised treatment based on the patient's condition. As such, the charges associated with each patient encounter should reflect the random nature of illness. Benford's law asserts that the occurrence of naturally occurring numbers, health care treatment charges in our case, conform to logarithmic based distributions (Benford (1938)).

There are however identifiable factors that influence the cost of treatment other than the condition being treated. Some examples of non-clinical factors that influence the cost of treatment include defensive medicine (Kessler and McClellan (1996), Studdert *et al.* (2005), Sloan and Shadle (2009), and Hermer and Brody (2010)), the circumstance in which physicians order excessive tests and procedures when faced with an increased threat of lawsuits. On the other hand, those who are uninsured and to a lesser extent those who are insured, forego needed treatment that is prohibitively expensive (Hadley *et al.* (1991)). We identify inefficiencies in health care, showing non-clinical factors are most prevalent in high severity encounters.

Supplier-induced demand (SID) (Richardson and Peacock (2006) and van Dijk *et al.* (2013)) occurs when patients, facing severe asymmetric information regarding their true health status, shift their health care consumption preferences to those of the care provider. Technology-driven demand (TDD) (Okunade and Murthy (2002), Smith *et al.* (2009), and Chandra and Skinner (2012)) occurs because the increased utilisation of cutting-edge technology increases costs and extends life expectancy. Chandra and Skinner (2012) show that health benefits of additional procedures, which significantly contribute to the high cost of health care, converge towards zero.

These other factors also suggest that the cost of treatment does not necessarily reflect the severity of the health ailment. These additional costs become more prevalent with the level of treatment, and the appropriate method of treatment becomes unclear. For example, routine check-ups and minor ailments present providers with a reduced risk of lawsuits as compared to complex surgical procedures or life-threatening health conditions. As the prevalence of non-random, human intervention increases, number distributions of total charges will increasingly deviate from the distribution predicted by Benford's law. These observations lead us to test that as health ailments become more severe, as measured by the cost of treatment, the impact of non-ailment-related factors strengthens. To our knowledge, we are the first to examine this question.

The rest of the paper is organised as follows. In Section I we describe Benford's law and present our hypotheses. Section II presents our data, methodology, and results. In Section III we conclude.

1. Benford's law and hypothesis development

1.1 Benford's law

Discovered in 1881 by Simon Newcomb, forgotten, and subsequently rediscovered and popularised by Frank Benford (1938), Benford's law asserts that digits of naturally occurring numbers conform to distributions based on logarithms. For example, the occurrence of the digit 1 as the first digit is expected to be [(LOG10(1 + (1/1)) ≈] 0.3010[1]. Figure 1 depicts the expected distribution of Benford's law. If a distribution of numbers is naturally occurring, then it is expected to follow the Benford distribution. Numbers such as city populations, levels of brightness in nature, and the size of lakes and ponds were all found to follow the expected logarithmic distribution (Benford (1938)).

Benford's law has been used to identify non-random human behaviour to investigate tax evasion (Nigrini (1996)), crypto-currency manipulation (McInish and Miller (working paper), and Covid-19 test results (Koch and Okamura (working paper) and Lee *et al.* (2020)). The key aspect of our research that is in common with these papers is detecting non-random behaviour in an attempt to protect the public from being harmed by inefficient health care delivery[2]. Specifically, deviations from Benford's law have indicated non-random, human intervention. In our study we evaluate total charges for health care consumption. If treatment is perfectly aligned with the medical ailment, and not influenced by non-random human behaviour, total charges for health care encounters are expected to follow Benford's law. While this figure is not the amount that is ultimately exchanged, it does represent the rawest indication of illness[3]. Avoiding public (and private) harm through efficient health care is an important endeavour for policy makers, academics, and the public. Using the Law of Anomalous Numbers, Benford's law, we analyse our distribution of encounter charges to determine if they are naturally occurring or if they are affected by non-clinical intervention. To our knowledge, we are the first to apply Benford's law to determine the non-constant effects of non-clinical factors on the cost of health care.

1.2 Hypotheses

Pointing to Winnie Langley, 'Britain's oldest smoker', Smith (2011) states that 'in general, Epidemiologists do a rather poor job of predicting who is and who is not going to develop a disease.'

[1]Benford's law predicts a positively skewed distribution of first- and second-digit numbers for naturally occurring distributions.

[2]This is not an indication of purposeful manipulation or nefarious activity on the part of the providers or patients.

[3]Health care providers usually collect a fraction of charges based on pre-negotiated prices with insurance companies. Also, most providers are unable to turn away indigent patients for lack of ability to pay meaning that some non-trivial proportion of encounters go unpaid. Under these conditions the amount of money that is ultimately collected deviates in non-random amounts from what is charged. This deviation makes the money collected for services unsuited for tests using Benford's law. The amount charged reflects the least amount of post-care adjustments to hospital charges available and so the best data for Benford's law tests.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

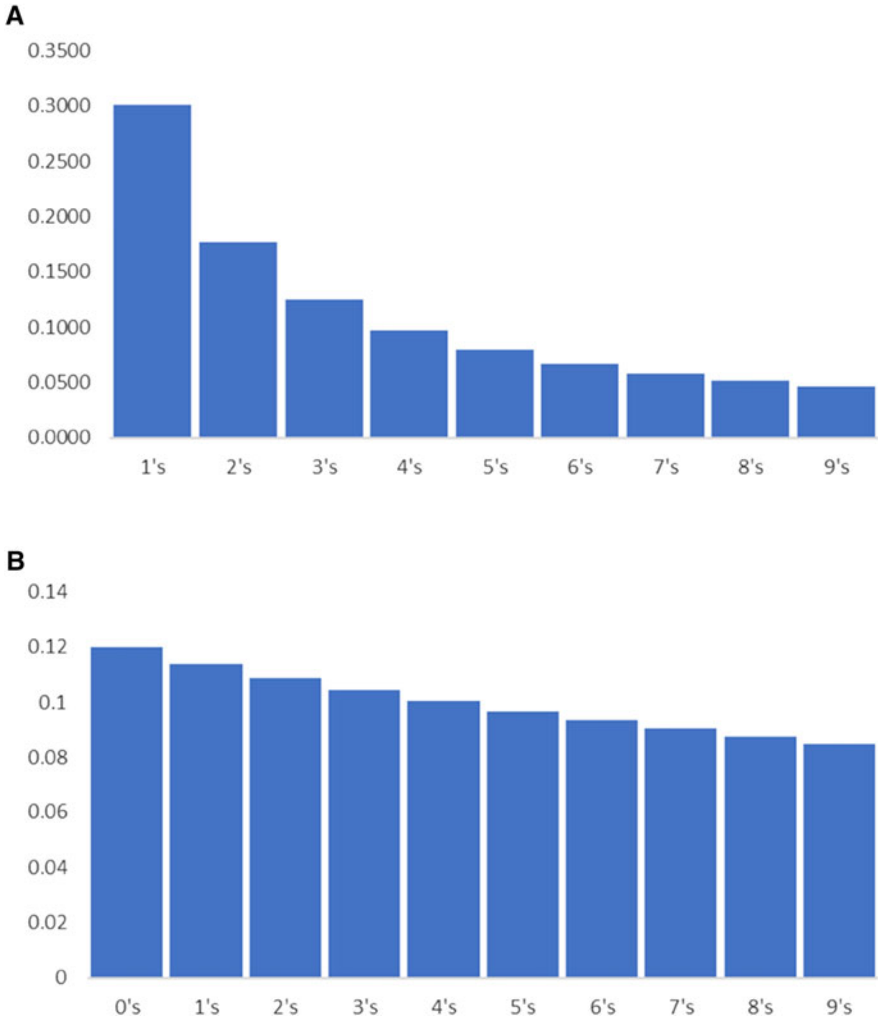

Figure 1. Panel A: Expected distribution of first digits. Panel B: Expected distribution of second digits. Panel A of Figure 1 is a histogram of the expected distribution of first digits for a given magnitude of 10 according to Benford's law. The vertical axis is the expected proportion of each of the possible first digits. The proportions sum to 1. Panel B is a histogram of the expected distribution of second digits according to Benford's law.

Smith argues that we should embrace the randomness of those with diseases. Yulmetyev *et al.* (2005) model the chaos and randomness in human health as well as the effectiveness of treatment.

Observable factors, including a patients age, gender, and geographical location may influence the extent of health care needs. Provider and payer factors including hospital size, urban setting, and payer type may also influence the availability of treatment. Non-observable factors such as genetic markers, immunocompetence, and illness potency certainly contribute to an individual's health experience. The set of unobserved factors means that general population predictions can be dubious for an individual. In embracing the randomness of individual predictions as argued by Smith (2011), our first hypothesis becomes:

Hypothesis 1: The severity of an individual's illness, as measured by the cost of treatment is unpredictable.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

If the cost of treatment reflects the severity of an individual's illness, then health care treatment would reflect the random nature of disease. However, certain market factors incentivise non-random human behaviour to drive treatment costs. Bell (1984) discusses New York State medical malpractice reform laws that cap payments to injured patients. He argues that consumers should only care about the legislation if the legislation results in an increase in medically unsafe behaviour, but not based on reduced charges to patients. Kessler and McClellan (1996) identify the behaviour of defensive medicine, wherein physicians order or perform costly treatments with minimal beneficial effect to avoid the financial and non-financial consequences of a malpractice lawsuit. Interestingly, these additional procedures expose providers to more malpractice risk and increases the provider's implicit marginal cost per procedure, possibly reducing utilisation of 'extra' procedures (Chandra and Skinner (2012), Currie and MacLeod (2008), and Baicker *et al.* (2007)).

McFarland *et al.* (working paper) tests the covariant relation between health care claim frequency and severity. They find that over the entire distribution of health care claims, the relation between frequency and severity is positive, but heterogeneous. As a patient access health care more frequently, the cost of each health care encounter increases. Patients with the most severe health ailments have more exposure defensive medicine, SID, and TDD, through both opportunity and cost[4]. These observations lead us to our second hypothesis.

Hypothesis 2: Health care costs deviate more from Benford's law as severity (costs) increase.

Newhouse (1992) claims that a consequence of 'too much' health insurance is 'too much' technological change. He finds that having health insurance leads to patients receiving extra treatment with advanced medicine that they otherwise would not receive. This is because health insurance drives the marginal price of health care to near zero. This is especially pronounced for socially funded health care coverage that is costless, or nearly costless to the insured. This circumstance highlights the presence of moral hazard. Privately insured patients are typically responsible for co-pays and deductibles, possibly mitigating the wasteful nature of 'too much' insurance.

We further test insurance coverage by type of payer to determine the prevalence of non-clinical factors among privately insured patients compared to government insured patients. Two factors incentivise monitoring by private insurers but not government insurers. First, individuals and employers must cover the cost of their private health insurance and second, private insurers seek to earn a profit. Further, Pauly (2000) argues that spending effects increase when insurance shields the consumer from financial responsibility, as is the case with many government-funded insurance. These observations lead us to our third hypothesis.

H3: Health care costs become less random when expenses are covered by government-funded insurance.

Uninsured individuals face the unique non-clinical factor of health care consumption of complete risk acceptance. Being fully and personally financially responsible for health care consumption alters uninsured individual's consumption choices. Hadley *et al.* (1991) find that uninsured individuals forego expensive treatment far more often than insured individuals. The increasing costs of health care coupled with the lack of risk sharing through insurance often prices uninsured individuals out of the market for health care services. Our fourth hypothesis is:

[4]Inefficient health care delivery may lead to excessive costs associated with the same encounter or lead to additional future encounters that are unnecessary or both. It is also possible that inefficient care by means of insufficient care leads to undo reductions in charges for either the encounter in question, a needed but foregone future encounter or both. In any case the distributions predicted by Benford's law would be violated.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

H4: Uninsured patient charges deviate from Benford's law across the entire severity distribution.

2. Data and methodology

2.1 Data

Our primary data source is the electronic health records Health Facts EMR dataset, made available through the Center for Biomedical Informatics at the University of Tennessee Health Science Center, UTHSC. The Health Facts dataset includes over 49 million distinct patients with more than 290 million patient encounters from 2000 through 2015. Data in Health Facts are extracted directly from the EMR from hospitals in which Cerner has a data use agreement. Cerner Corporation has established Health Insurance Portability and Accountability Act-compliant operating policies to establish de-identification for Health Facts.

Each encounter begins upon admission and ends at discharge. The total charges for an encounter, our main variable of interest, are the summation of all provider-related charges for that encounter. Charges for outpatient prescriptions, when written by a primary care provider, but not filled during a clinical visit, are excluded from total charges. However, inpatient prescriptions administered by the provider during the encounter are included in total charges[5].

It is preferable that the data set covers multiple magnitudes (1s, 10s, 100s), covers a full range of magnitude[6], and that the data are not averaged. It is also critical that the numbers are not rounded or have minimums[7] or maximums. Our data complies with these criteria.[8] Distributions that are expected to follow Benford's law include transactions-level data (e.g. sales, trade size), numbers that result from a combination of numbers – quantity × price. Data for which the mean is greater than the median are also more likely to follow Benford's law. McFarland *et al.* (working paper) find that the distribution of individual health care costs is positively skewed (Figure 2).

Using patient billing data, we evaluate the scope of total encounter expenditures by reviewing admission sources and discharge dispositions to identify patients who are expected to have additional encounters. This approach facilitates our developing parameters for estimating and capturing health care expenses. We apply our filters to ensure that we include only those patients for whom we have most claim data[9]. This leaves us with over 59 million encounters. For much of our study, we further limit our observations to encounter with a minimum charge of \$100 and not exceeding \$1,000,000. This limitation allows us to group encounters by severity while maintaining full magnitudes of 10 in the first digit. Given our filters, our sample covers the

[5]Regarding billing data, the total charges variable represents hospital invoice charges before receiving any deductions in received/reimbursed payments. Payments remitted by insurance companies to health care providers are reduced by agreements between parties. We recognize the potential overstatement nature of the data; however, our study focuses on charged amounts as a reflection of the random nature of illness.

[6]Temperature (in Fahrenheit) ranging from 30 to 95° for example.

[7]We remove from our final sample charges that range from \$0 to \$99.99. We expect that there is an economic minimum amount charged, even if not formally stated, for the most minor of encounters providers still need to price in overhead costs, labour costs, and administrative costs. These considerations would lead us to expect that the lowest range (\$0–\$99.99) would reasonably deviate from Benford's law dramatically for reasons not related to inefficiency in the sense that our paper is addressing.

[8]The occurrence rate of an encounter in excess of \$100,000 is of course lower than the occurrence rate of an encounter between \$1000 and \$9999.99. However, *of the encounters that are in excess of \$100,000* we should still see a distribution of first and second digits according to Benford's law because these encounters are none the less naturally occurring, positively skewed, and cover a full magnitude of 10. In other words, when using raw counts, the slope of the first (second) digit count distribution remains constant across each price bucket even if the intercept is decreasing. Consistent with prior literature we report percentages rather than raw counts.

[9]We eliminate observations that have an admission (discharge) status of transferred in (transferred out) but no preceding (proceeding) encounters.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

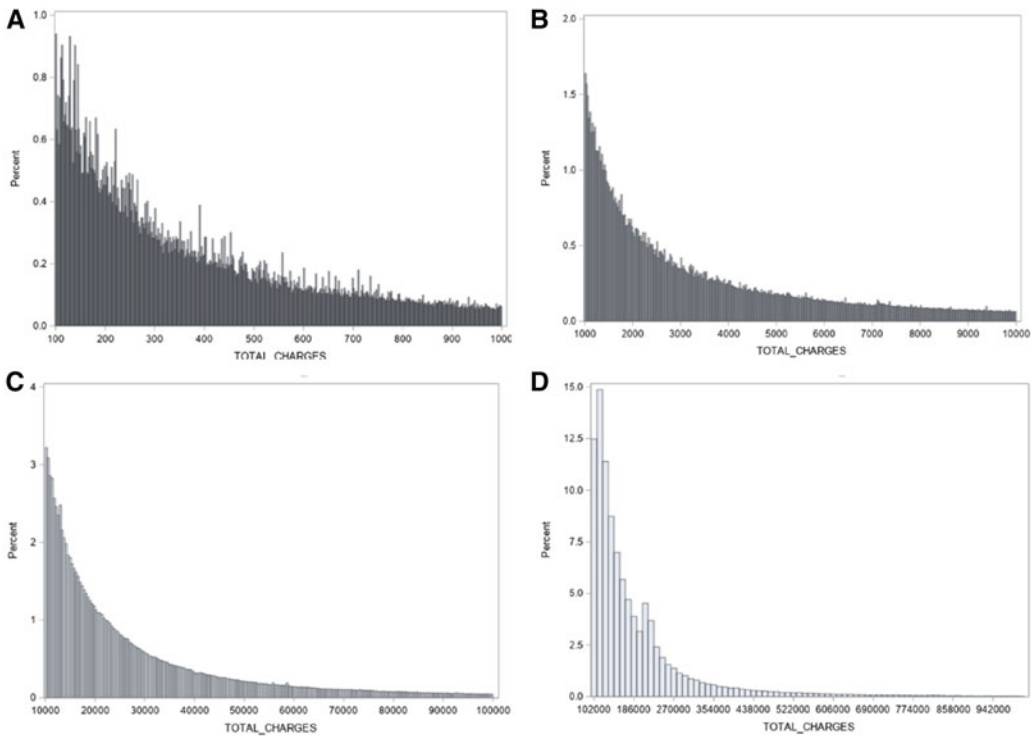

Figure 2. Panel A: Price bucket 1, Panel B: Price bucket 2, Panel C: Price bucket 3, Panel D: Price bucket 4. Figure 2 is a set of histograms showing the distribution of total charges for each price bucket.

years 2000 through 2014 and includes over 47 million encounters. Table 1 reports descriptive statistics regarding patients, encounters, encounter charges, and length of stay.

We report our descriptive statistics both as a single sample and segmented by charged amount. Encounters with charges between \$100 and \$999.99 (charges = 1), \$1000 and \$9999.99 (charges = 2), \$10,000 and \$99,999.99 (charges = 3), and \$100,000 and \$999,999.99 (charges = 4) are grouped together. We observe a negative relation between encounter severity (measured by encounter charges) and the number of encounters. Health care charges, like most insurable risks, are skewed distributions wherein relatively few people experience extremely high health care costs. Notwithstanding, we find a significant number of encounters at all severity levels, including nearly 200,000 encounters with charges in excess of \$100,000. Our sample includes more female encounters (36 million) than male encounters (23 million). Most of our encounters are patients that are aged 18–65 (34 million) while the average encounter charge appears to increase with the age of the patient. With the notable exception of research-based encounters, our sample includes an even mix of payer types with over 19 million government payer encounters, 12 million commercial payers, and 3 million self-payers.

2.2 Methodology

Following Nigrini and Mittermaier (1997) we use three tests: first-digits, second-digits, and first-two digits. We also follow Drake and Nigrini (2000) by calculating the mean of absolute deviations (MAD) to use as a way to assess conformity to the expected distribution. A naïve person choosing numbers at random would most likely guess a distribution would be uniform with 11% occurring for each digit 1–9. However, Benford's law states that for a group of natural

Table 1. Descriptive Statistics

	N	Min charges	Mean charges	Max charges	Min LOS	Mean LOS	Max LOS
All	59,144,715	0.01	$3,579	$7,901,714	<1	1.40	2,744
Charges = 1	29,298,060	100	366	999	<1	1.02	94
Charges = 2	13,656,692	1,000	3,330	9999	<1	1.36	712
Charges = 3	4,235,846	10,000	27,350	99,999	<1	4.36	2,744
Charges = 4	197,802	100,000	188,062	999,999	<1	21.10	1,431
Female = 1	36,112,412	0.01	3,186	7,901,714	<1	1.37	2,744
Female = 0	22,996,124	0.01	4,198	6,608,907	<1	1.47	1,355
Age <18	7,749,714	0.01	2,971	6,160,271	<1	1.42	1,087
Age 18-65	33,970,146	0.01	3,220	4,984,151	<1	1.32	1,027
Age >65	17,424,855	0.01	4,549	7,901,714	<1	1.57	2,744
Government	19,789,954	0.01	3,752	6,608,907	<1	1.46	1,987
Commercial	12,429,041	0.01	2,766	7,901,714	<1	1.24	1,087
Self	3,038,432	0.01	2,948	2,110,052	<1	1.31	2,744
Research	30,325	5	2,689	1,403,146	<1	1.23	126
Other	23,856,963	0.01	3,940	6,160,271	<1	1.46	1,355

We examine the cost of health care encounters for patients across the U.S. during the years 2000 through 2014 from the Health Facts EMR data. We report the distribution of encounter *charges* and encounter length of stay (*LOS*), measured in days, for all patients. We present our results for the entire sample (*All*) and classified by the charged amount (*Charges*), the patient gender (*Female*), patient age (*Age*), and payer type. For each category we report *minimum*, *mean*, and *maximum charges* and *LOS*.

occurring numbers the distribution of first digits occurs based on logarithmic properties as follows: 1s 0.3010, 2s 0.1761, 3s 0.1249, 4s 0.0969, 5s 0.0792, 6s 0.0669, 7s 0.0580, 8s 0.0512, and 9s 0.0458. In addition to the standard test of counts of first digits, we calculate the mean average deviations (MAD), which is called a reasonableness test by Drake and Nigrini (2000). To compute the MAD, we average the absolute value of deviations and divide by 9. For first digits, a MAD of 0.000 ± 0.004 indicates close conformity, 0.004 ± 0.008 acceptable conformity, 0.008 ± 0.012 marginally acceptable conformity, and a MAD greater than 0.012 nonconformity.

Smith (2011) acknowledges that epidemiologists do a poor job of determining who is going to get sick with what ailment and when. This is because of the readily acceptable fact that illness affects individuals randomly. Additionally, the causes and complications of diseases may be determined by variables not readily observable, either ax-ante or ex-post. We begin our analysis by identifying patient, facility, payer, or regional factors that contribute to the cost of a health care encounter. Master diagnostic codes (MDCs) are nationally recognised standard groupings that correspond to single organ system ailments or medical specialties e.g. respiratory. We expect that overall encounter cost levels vary by MDC group. However, each MDC group encompasses a wide range of encounter types, from low-cost preventative care to the severe emergent encounters. We begin by regressing encounter charges on MDC controls and estimate our first regression model as

$$EC = \beta_1 + X_{\text{MDC}} + \varepsilon \tag{1}$$

where EC is an abbreviation of encounter charges and X_{MDC} is a vector of MDC dummy variables. In model 2 we include patient and calendar year descriptive variables. *Age* is the patient's

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

age in years at the time of admission. *Year* is the calendar year at admission to control for rising health care costs over time. We also include vectors of dummy variables for *gender (G)*, *race (R)*, *marital status (M)*, and *US census location (L)* of the patient.

$$EC = \beta_1 + \beta_2 Age + \beta_3 Year + X_{\mathrm{MDC}} + G_{\mathrm{Gender}} + R_{\mathrm{Race}} + M_{\mathrm{Marital}} + L_{\mathrm{location}} + \varepsilon \tag{2}$$

In model 3 we retain our previous control variables and include the treating facility variables *urban*, a dummy variable equal to 1 for all urban providers, *size* (the size of the facility based on licensed bed count), and *teaching*, a dummy variable equal to 1 for all teaching hospitals.

$$\begin{aligned} EC = {} & \beta_1 + \beta_2 Age + \beta_3 Year + \beta_4 Urban + \beta_5 Size + \beta_6 Teaching + X_{\mathrm{MDC}} \\ & + G_{\mathrm{Gender}} + R_{\mathrm{Race}} + M_{\mathrm{Marital}} + L_{\mathrm{Location}} + \varepsilon \end{aligned} \tag{3}$$

In model 4 we expand our control variables to include time-of-week and year. *Weekday* is a dummy variable equal to 1 for all admission that occur Monday through Friday. *Holiday* is a dummy variable equal to 1 for admissions that occur on a nationally recognised holiday. We also include a vector of monthly dummy variables (*Mt*) to account for seasonal effects.

$$\begin{aligned} EC = {} & \beta_1 + \beta_2 Age + \beta_3 Year + \beta_4 Urban + \beta_5 Size + \beta_6 Teaching + \beta_7 Weekday \\ & + \beta_8 Holiday + X_{\mathrm{MDC}} + G_{\mathrm{Gender}} + R_{\mathrm{Race}} + M_{\mathrm{Marital}} + L_{\mathrm{Location}} \\ & + Mt_{\mathrm{Month}} + \varepsilon \end{aligned} \tag{4}$$

Finally, our last model includes the aforementioned control variables and *price bucket*.

$$\begin{aligned} EC = {} & \beta_1 + \beta_2 Age + \beta_3 Year + \beta_4 Urban + \beta_5 Size + \beta_6 Teaching + \beta_7 Weekday \\ & + \beta_8 Holiday + \beta_9 Price\ Bucket + X_{\mathrm{MDC}} + G_{\mathrm{Gender}} + R_{\mathrm{Race}} \\ & + M_{\mathrm{Marital}} + L_{\mathrm{Location}} + Mt_{\mathrm{month}} + \varepsilon \end{aligned} \tag{5}$$

2.3 Empirical results

2.3.1 Does encounter severity reflect the random nature of illness severity?

In Table 2 we subdivide our sample into four price buckets, the first for all encounters that incur costs of at least \$100[10] and not more than \$999.99. The second price bucket includes encounters with charges ranging from \$1000 to \$9999.99. Price bucket 3 includes all encounters with charges between \$10,000 and \$99,999.99, and the final price bucket includes charges of at least \$100,000 but not exceeding \$999,999.99. We find that while many of our control variables are statistically significant, and in some cases economically significant as well, the best any of these models can do is return an r-squared of less than 0.03. In all cases except for model 3 the largest coefficient in terms of magnitude is the intercept term. These results support our first hypothesis that the severity of illness, and the associated costs are random at the individual level.

2.3.2 Non-clinical factors

In a frictionless environment the cost of a health care encounter should reflect the severity of the health care ailment. However, it is well understood that the real world is not frictionless. Within

[10]We exclude observations with charges less than \$100 as well as observations with charges in excess of \$1,000,000 as they are a minor part of our sample and would distort results as they do not fully cover the \$10–\$100 or \$1,000,000–\$10,000,000 ranges of magnitude.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

Table 2. OLS estimated encounter expense.

	Model 1	Model 2	Model 3	Model 4	Model 5
R^2	0.0169	0.0234	0.0248	0.0261	0.0261
Intercept	5221***	3000***	−26.7	1786***	1680***
Age		27***	27***	29***	29***
Year		52***	70***	75***	75***
Urban			1230***	1175***	1173***
Size			435***	443***	441***
Teaching			−59***	−14**	−10
Weekday				−2125***	−2133***
Holiday				808***	813***
Price bucket					20***
MDC	Yes	Yes	Yes	Yes	Yes
Gender		Yes	Yes	Yes	Yes
Race		Yes	Yes	Yes	Yes
Marital		Yes	Yes	Yes	Yes
Location		Yes	Yes	Yes	Yes
Month				Yes	Yes

We estimate OLS regressions to determine factors that predict individual encounter charges from the Health Facts EMR data. *Age* is the patient's age in years, *Year* is the calendar year, *Urban* is a dummy variable equal to 1 for patients that access an urban provider. *Size* is the bed size of the providing hospital. *Teaching* is a dummy variable equal to 1 if the hospital is a teaching hospital. *Weekend* and *Holiday* are dummy variables indicating the day of admission. We include fixed effects for MDC code, gender, race, marital status, US census location, and month. In model one we include only controls for MDC. In model two we also include *age* and *year* controls. Model three includes provider variables and in model four we also include *weekday* and *holiday* control variables. Final, in model five we include a control for price bucket. *, **, and *** indicate significance at the 10%, 5%, and 1% significance levels, respectively.

the setting of health care, there are some readily identifiable frictions, non-clinical factors that directly affect the cost of health care. Defensive medicine, SID, and TID are some of frictions. The magnitude and presence of these factors are difficult, if not impossible to identify at the encounter level. Additional unnecessary procedures, examinations, or diagnostic tests are justified by individual provider judgements or out of an abundance of diagnostic scepticism as opposed to nefarious motives. Many studies identify the presence of non-clinical cost factors by observing changes in expenditure before and after legislative developments (Kessler and McClellan (1996), Sloan and Shadle (2009), adoption of technologically advanced treatments (Chandra and Skinner (2012), and R&D spending. We execute an alternative research design to identify the presence of non-clinical factors in health care charges by applying Benford's law to encounter charges. This methodology does not distinguish between different non-clinical factors but does provide insight into the magnitude of the non-clinical factors at varying levels of encounter severity. This identification strategy is important to policy makers, medical practitioners and providers because it allows them to focus cost efficiency efforts on the relatively few encounters with the most inefficiencies. Benford's law applies nicely to number distributions that are naturally occurring, cover multiple magnitudes, and are positively skewed, as is the case with our data. Encounter level health care costs should meet these three criteria well if non-clinical charges are not present. We can therefore attribute most non-conformity to non-clinical factors.

We segment our sample of encounters into subsamples based on the charges for each encounter. We identify four subsamples of encounters consistent with the *price bucket* variable defined in Table 2. Segmenting the encounters this way provides at least two important benefits to our study. The first is that each *price bucket* contains a range that begins with a lowest possible first-digit number equal to 1 and ends with a highest possible first-digit number equal to 9. Furthermore, the distribution spans only a single magnitude of 10 for each distribution, meaning each total-charges bucket provides the same a priori probability to all digits 1 through 9 of being the first digit. The second important benefit of partitioning our sample is that it allows us to evaluate the conformity of Benford's law across different encounter severities[11]. We expect that the non-clinical factors are more prevalent for more severe encounters. For example, defensive medicine occurs as a deterrent to malpractice lawsuits. However, the risk of a malpractice lawsuit is less when the severity of an illness is small. Therefore, physicians will be more likely to practice defensive medicine, and in greater quantities, as the encounter severity increases[12]. Because of this we expect that the distribution of total charges deviates in greater degree as total-charges increase. The same can be said for SID and TDD factors.

We employ the count test to examine first-digits (leading-digits). Table 3 presents our results. We find that at the MAD for the first bucket of charges (0.010) shows a marginally acceptable conformity to Benford's law. However, for the second (0.023), third (0.049), and fourth (0.092) buckets the MAD is greater than 0.012 indicating nonconformity. As expected, the MAD increases with the level of total-charges. Interestingly, we find consistent deviations from the predicted probabilities at the price bucket boundaries. 1-as-the-leading-digit is consistently over-represented while 9 is consistently under-represented. An additional possible explanation for this finding is hospital pricing strategies (Krishnan (2001), Sutherland (2015)). If providers are strategically pricing their services, then either prices are being strategically raised or lowered. If prices are being raised, our boundary observations show that services near the high end of a price range are being raised sufficiently to move those services into the next price bucket. This would cause an increase in 1-as-the-leading-digit occurrences and a decrease in 9-as-the-leading-digit occurrences, consistent with our observation. Alternatively, if prices are strategically lowered, then we would find the opposite result. In this case, strategic pricing strategies are muting the effect of other non-clinical factors affecting health care costs, and actual inefficiencies are more severe than we can identify. Hospital pricing strategies are beyond the scope of our research but present a compelling case for further research. This finding supports our second hypothesis that encounter charges deviate from Benford's law as the total-charges increase.

According to Benford's law the second-digits expected probabilities for each number are as follows: 0s 0.1197, 1s 0.1139, 2s, 0.1088, 3s 0.1043, 4s 0.1003, 5s 0.0967, 6s 0.0934, 7s 0.0904, 8s 0.0876, 9s 0.085. This is much closer to uniform, but a dataset of uniform second-digits is statistically different from the expected distribution. We repeat our previous procedure on the second-digits, but the MAD thresholds are slightly different. For second digits, a MAD of 0.000 ± 0.008 indicates close conformity, 0.008 ± 0.012 acceptable conformity, 0.016 ± 0.016 marginally acceptable conformity, and a MAD greater than 0.016 nonconformity (Table 4).

Examining the distribution of second digits we find that total charges in our first two buckets, \$100–\$999.99 and \$1000–\$9999.99 closely conform to the Benford distribution. For our third bucket, encounters with total charges between \$10,000 and \$99,999.99 we see acceptable conformity. For our bucket with our most expensive encounters, like our test of first digits we find non-conformity indicating some type of thoughtful intervention. The consistent decline in conformity to the expected logarithmic distribution supports our second hypothesis and

[11]By setting a minimum of \$100 for the first price bucket and a maximum of \$999,999.99 for the fourth price bucket we lose 11,756,315 observation. As noted previously, charges below \$100 will have a natural minimum that is greater than 1 meaning this range will not cover a full magnitude of 10. The maximum charges in our sample are under \$8 million, so a price bucket in excess of \$1 million also will not cover a full magnitude of 10.

[12]TDD is similar in that more advanced technology may be applied in more severe illness but not in less severe illness. Other non-clinical factors follow the same pattern.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

Table 3. Distribution of first-digits

		$100–$999		$1000–$9999	
Digit	Expected per cent	Frequency	Per cent	Frequency	Per cent
1	30.01	9,127,610	31.15	5,336,612	39.08
2	17.61	5,846,948	19.96	2,595,328	19.0
3	12.49	3,866,295	13.20	1,656,229	12.13
4	9.69	2,993,038	10.22	1,146,780	8.40
5	7.92	2,201,815	7.52	873,107	6.39
6	6.70	1,730,524	5.91	663,187	4.86
7	5.80	1,469,020	5.01	558,501	4.09
8	5.11	1,125,703	3.84	447,934	3.28
9	4.58	937,107	3.20	379,014	2.78
MAD			0.010[3]		0.023[4]
		$10,000–$99,999		$100,000–$999,999	
Digit	Expected per cent	Frequency	Per cent	Frequency	Per cent
1	30.1	2,068,483	48.83	140,321	70.94
2	17.61	888,166	20.97	36,670	18.54
3	12.49	470,192	11.10	9,869	4.99
4	9.69	281,806	6.65	4,692	2.37
5	7.92	185,936	4.39	2,563	1.30
6	6.70	127,375	3.01	1,536	0.78
7	5.80	93,409	2.21	991	0.50
8	5.11	68,577	1.62	702	0.35
9	4.58	51,902	1.23	458	0.23
MAD			0.049[4]		0.092[4]

Using the Health Facts EMR data, we partition the observed encounters into four buckets based on total charges. The first bucket includes all observations with total charges of at least $100 and less than $1000. The second bucket includes total charges of at least $1000 and less than $10,000. The third bucket includes all charges of at least $10,000 and less than $100,000. The fourth bucket includes all charges of at least $100,000 and less than $1,000,000. We report the frequency of each digit (1 through 9) occurring as the first digit as a frequency as well as a per cent of the total in the bucket. Mean absolute deviations (MAD) are presented with 1 for indicating close conformity, 2 acceptable conformity, 3 marginally acceptable conformity, and 4 non-conformity.

indicates that more non-clinical factors such as defensive medicine, SID, and TDD are present as the total charges of a patient's encounter increase.

2.3.3 Insurance and payer type

Health care in the US is predominantly financed through insurance. The risk sharing characteristics of insurance creates some separation between the consumer of health care and the payer. When the government, at any level, is the payer, this separation is amplified. A relatively small fraction of patients, however, are self-insured or otherwise pay for health care services directly. These important differences in the degree of separation between consumer and payer may lead to different applications or magnitudes of non-clinical cost factors. For example, Newhouse (1992) argues that 'too much' insurance may lead to 'too much' technological change.

Government insured patients are the consumers of health care that are furthest removed from the costs of health care. This is due to government-funded health insurance is generally made

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

Table 4. Distribution of second-digits

		$100–$999		$1000–$9999	
Digit	Expected per cent	Frequency	Per cent	Frequency	Per cent
0	11.96	3,355,973	11.45	1,808,363	13.24
1	11.39	3,373,298	11.51	1,639,321	12.00
2	10.90	3,243,949	11.07	1,565,187	11.46
3	10.43	3,099,256	10.58	1,469,972	10.76
4	10.03	3,017,218	10.30	1,359,007	9.95
5	9.67	2,906,994	9.92	1,293,371	9.47
6	9.34	2,702,606	9.22	1,212,116	8.88
7	9.03	2,610,884	8.91	1,168,263	8.55
8	8.76	2,540,908	8.67	1,091,389	7.99
9	8.50	2,446,974	8.35	1,049,703	7.69
MAD		0.0020[1]		0.0053[1]	
		$10,000–$99,999		$100,000–$999,999	
Digit	Expected per cent	Frequency	Per cent	Frequency	Per cent
0	11.96	598,631	14.13	37,513	18.96
1	11.39	548,983	12.96	35,739	18.07
2	10.90	509,450	12.03	26,992	13.65
3	10.43	457,407	10.80	21,450	10.84
4	10.03	420,184	9.92	17,855	9.03
5	9.67	387,897	9.16	15,171	7.67
6	9.34	364,546	8.61	13,042	6.59
7	9.03	336,851	7.95	11,457	5.79
8	8.76	316,134	7.46	9,785	4.95
9	8.50	295,763	6.98	8,798	4.45
MAD		0.0099[2]			0.0329[4]

Using the Health Facts EMR data, we partition the observed encounters into four buckets based on total charges. The first bucket includes all observations with total charges of at least $100 and less than $1000. The second bucket includes total charges of at least $1000 and less than $10,000. The third bucket includes all charges of at least $10,000 and less than $100,000. The fourth bucket includes all charges of at least $100,000 and less than $1,000,000. We report the frequency of each digit (0 through 9) occurring as the second digit as a frequency as well as a per cent of the total in the bucket. Mean absolute deviations (MAD) are presented with 1 for indicating close conformity, 2 acceptable conformity, 3 marginally acceptable conformity, and 4 non-conformity.

available at little or no cost to those who cannot otherwise afford private health insurance. Being so far removed from the cost of health care means that however minimal the benefits of additional treatment, the patient will likely accept treatment since they are not responsible for the costs. The patient's financial incentives for partaking in defensive medicine, SID, or TDD are aligned with the provider's and suppliers' non-clinical incentives to provide such care. On the other hand, those who are privately insured are financially responsible for co-pays, deductibles, or a portion of treatment costs. These financial responsibilities detach the patient's clinical incentives from provider's non-clinical incentives. Self-insured patients represent the extreme disconnect between clinical and non-clinical incentives to the point that self-insured patients may

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

Table 5. Distribution of first- and second-digits by payer

Panel A: First digits		Severity bucket			
Payer	*N*	1	2	3	4
Government	15,915,860	0.012[3]	0.024[4]	0.045[4]	0.101[4]
Private	9,927,441	0.009[3]	0.028[4]	0.056[4]	0.099[4]
Self	2,461,244	0.013[4]	0.034[4]	0.055[4]	0.104[4]
Panel B: Second digits					
		Severity bucket			
Payer	*N*	1	2	3	4
Government	15,915,860	0.003[1]	0.006[1]	0.009[2]	0.033[4]
Private	9,927,441	0.002[1]	0.007[1]	0.013[3]	0.034[4]
Self	2,461,244	0.004[1]	0.008[1]	0.012[3]	0.035[4]

Using the Health Facts EMR data, we segment the observed encounters by payer type and partition each sub-sample into four buckets based on total charges. The first bucket includes all observations with total charges of at least $100 and less than $1000. The second bucket includes total charges of at least $1000 and less than $10,000. The third bucket includes all charges of at least $10,000 and less than $100,000. The fourth bucket includes all charges of at least $100,000 and less than $1,000,000. We report the mean absolute deviations (MAD) and the corresponding level of conformity. 1 indicating close conformity, 2 acceptable conformity, 3 marginally acceptable conformity, and 4 non-conformity.

forego clinically prescribed treatments if the costs are prohibitively expensive or if the costs outweigh the perceived benefits (Hadley *et al.* (1991)).

Patient and provider incentives are not the only means by which payer types may result in different results to our study. Private insurers operate to earn a profit, whereas government insurance programmes do not. The profit motive creates the incentive for private insurers to monitor and prevent non-clinical cost factors. We test these observations by segmenting our sample into three subsamples based on payer type. The first sub-sample comprises those patients with government-funded health insurance. Second, private health insurance and lastly self-insured patients. We remove from these subsamples of payers' patient encounters whose payment source is unidentified or research based. We calculate the distribution of first- and second-digit numbers and calculate the corresponding MAD values. Table 5 reports our results. Both government (0.012) and privately insured (0.009) patients show marginally acceptable conformity for the lowest severity bucket and nonconformity for all other severity buckets. While the MAD value for the lowest severity privately insured patients (0.009) is less than the MAD value for the lowest severity government insured patients (0.0112), in buckets 2 and 3 the rank order reverses. This finding is contrary to our third hypothesis that charges become less random when health care is government funded. The inability to effectively monitor providers, either due to cost-based reimbursement or loss-estimation difficulties (Pauly (2000)) are likely contributors to this finding.

When examining the count of first digits we also observe that in all cases self-insured patients do not conform to Benford's law. MAD values for self-insured patients are 0.013, 0.034, 0.055, and 0.104 for buckets 1 through 4, respectively. This suggests that non-clinical factors, both cost increasing (defensive medicine, SID, and TDD) and cost reducing (refusal of treatment), affect self-insured patients more than insured patients. This finding is in support of our fourth hypothesis that charges for self-insured patients deviate from Benford's law across the entire distribution of claim severity[13].

[13]We also perform our analysis based on emergency department visits and hospital size with similar results. See Appendix A for tabulated results.

https://doi.org/10.1017/S174413312400015X Published online by Cambridge University Press

3. Conclusion

The rising cost of health care has been proven to be caused in part by non-clinical factors. We find that as patients' total charges increase, their expenses begin to deviate more from the random nature of human illnesses. In our most severe, highest cost bucket, we find compelling evidence to reject conformity to Benford's law for both our test of the distribution of first digits and the distribution of second digits of medical charges, indicating that the most severe encounters contain the least random pricing. This finding is robust to payer type, provider type, and ED visits. A possible explanation for this is that as an illness becomes more severe the incentives for providing additional procedures also increase and the ability to monitor wasteful spending decreases. We suggest focusing cost reduction efforts on high severity encounters. The most severe encounters (buckets 3 and 4 combined) represent only 7.5% of the total encounters but over 72% of the dollars spent on health care treatment. Such efforts will focus the attention of providers, insurers, and regulators to effectively alleviate cost burdens caused by technological advancement or procedures that provide only marginal, if any, benefit. Finally, policies to encourage self-insured and uninsured patients to undertake needed procedures while monitoring against unnecessary procedures will promote greater health.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S174413312400015X.

Data availability. The data that support the findings of this study are available from the corresponding author upon request.

Acknowledgements. We thank the University of Tennessee Health Science Center and participants at the 2021 South Western Finance Association annual meeting. We also thank anonymous referees for the helpful

Financial support. The authors did not receive financial support from any organisation for the submitted work.

Competing interests. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

References

Baicker K, Fisher ES and Chandra A (2007) Malpractice liability costs and the practice of medicine in the Medicare Program. *Health Affairs* **26**, 841–52.

Bell PA (1984) Legislative intrusions into the common law of medical malpractice: thoughts about the deterrent effect of tort liability. *Syracuse L. Rev* **35**, 939.

Benford R (1938) Law of anomalous numbers. *American Philosophical Society* **4**, 551–572.

Chandra A and Skinner J (2012) Technology growth and expenditure growth in health care. *Journal of Economic Literature* **50**, 645–680.

Currie J and Bentley MacLeod W (2008) First do no harm? Tort reform and birth outcomes. *Quarterly Journal of Economics* **123**, 795–830.

Drake P and Nigrini M (2000) Computer assisted analytical procedures using Benford's law. *Journal of Accounting Education* **18**, 127–146.

Hadley J, Steinburg P and Feder J (1991) Comparison of uninsured and privately insured hospital patients' condition on admission, resource use, and outcome. *JAMA* **265**, 374–379.

Hermer LD and Brody H (2010) Defensive medicine, cost containment, and reform. *Journal of General Internal Medicine* **25**, 470–473.

Jackson P (1998) The impact of health insurance status on emergency room services. *Journal of Health and Social Policy* **14**, 1.

Kessler D and McClellan M (1996) Do doctors practice defensive medicine? *The Quarterly Journal of Economics* **111**, 353–390.

Koch C and Okamura H. Benford's law and COVID-19 reporting. Working paper. 4.

Krishnan R (2001) Market restructuring and pricing in the hospital industry. *Journal of Health Economics* **20**, 213–237.

Lee K, Han S and Jeong Y (2020) COVID-19, flattening the curve, and Benford's law. *Physica A: Statistical Mechanics and its Applications* **559**, 125090.

McInish T and Miller J. Detecting Fake Bitcoin Volume, working paper.

Newhouse JP (1992) Medical care costs: how much welfare loss? *Journal of Economic Perspectives* **6**, 3–21.

Nigrini MJ (1992) The detection of income tax evasion through an analysis of digital frequencies (*Thesis*). University of Cincinnati, Ohio.

Nigrini MJ (1996) A taxpayer compliance application of Benford's law. *Journal of American Taxation Associates* **18**, 72–91.
Nigrini MJ (2012) *Benford's Law Applications for Forensic Accounting, Auditing, and Fraud Detection*.
Nigrini MJ and Mittermaier LJ (1997) The use of Benford's Law as an aid in analytical procedures. *Auditing: A Journal of Practice and Theory* **16**, 52–67.
O'Brien, Stein M, Zierler S, Shapiro M, O'Sullivan P and Woolard R (1997) Use of the ED as a regular source of care: Associated factors beyond lack of health insurance. *Annals of Emerging Medicine* **3**, 286–291.
Okunad AA and Murthy VN (2002) Technology as a 'major driver' of health care costs: a cointegration analysis of the Newhouse conjecture. *Journal of Health Economics* **21**, 147–159.
Pauly M (2000) Insurance reimbursement. In Culyer AJ and Newhouse JP (eds), *Handbook of Health Economics*. Elsevier, pp. 537–560.
Richardson JR and Peacock SJ (2006) Supplier-induced demand: reconsidering the theories and new Australian evidence. *Applied Health Economics and Health Policy* **5**, 87–98.
Sloan FA and Shadle JH (2009) Is there empirical evidence for "defensive medicine"? A reassessment. *Journal of Health Economics* **28**, 481–491.
Smith GD (2011) Epidemiology, epigenetics and the 'Gloomy Prospect': embracing randomness in population health research and practice. *International Journal of Epidemiology* **40**, 537–562.
Smith S, Newhouse JP and Freeland MS (2009) Income, insurance, and technology: why does health spending outpace economic growth?. *Health Affairs* **28**, 1276–1284.
Studdert DM, Mello MM, Sage WM, DesRoches CM, Peugh J, Zapert K and Brennan TA (2005) Defensive medicine among high-risk specialist physicians in a volatile malpractice environment. *JAMA* **293**, 2609–2617.
Sutherland JM (2015) Pricing hospital care: Global budgets and marginal pricing strategies. *Health Policy* **119**, 1111–1118.
van Dijk, van den, Verheij RA, Spreeuwenberg P, Groenewegen PP and de Bakker (2013) Moral hazard and supplier-induced demand: empirical evidence in general practice. *Health Economics* **22**, 340–352.
Yulmetyev RM, Yulmetyeva D and Gafarov FM (2005) How chaosity and randomness control human health. *Physica A: Statistical Mechanics and its Applications* **354**, 404–414.

Cite this article: McFarland S and and Miller J (2024) The inefficient effects of non-clinical factors on health care costs. *Health Economics, Policy and Law* 1–15. https://doi.org/10.1017/S174413312400015X