Was Civil War Surgery Effective?

Matthew J. Baker*

December 29, 2016

Abstract

During the U. S. Civil War (1861-65) surgeons performed a vast number of surgical procedures. The efficacy of surgery has been continually debated since the war began, in part because of lack of evidence for the (in)effectiveness of surgery. I analyze data gathered by Dr. Edmund Andrews, a surgeon with the 1st Illinois Light Artillery. The data can be arranged as observational data on surgery and recovery, with controls for wound location and severity, and with instruments for surgery. Analysis of the data using bivariate probit and a switching regression suggests that surgery was effective, was applied selectively by surgeons, and increased the probability of survival with an average treatment effect of 0.06-0.25 points. Results also suggest that surgeons applied surgery selectively and in situations in which it was likely to be beneficial; among those receiving surgery, I find an average treatment effect of 0.25-0.28 points.

*Department of Economics, Hunter College and the Graduate Center, CUNY. Email: matthew.baker@hunter.cuny.edu. I would like to thank Jonathan Conning, Partha Deb, and Jessica Van Parys for helpful comments and suggestions.

I am faithful, I do not give out,
The fractur'd thigh, the knee, the wound in the abdomen,
These and more I dress with impassive hand (yet deep in my breast
a fire, a burning flame)
- Walt Whitman, *The Wound Dresser*

1 Introduction

The Civil War (1861-65) was a seminal event in United States History. As is often said, more American soldiers died in the Civil War than in all other American Wars combined, and the Civil War is to this day the most written-about event in United States history. The Civil War was also of great importance in the history of medicine. During the war, surgeons gathered a large amount of information on wound types and treatment. These efforts advanced medical and surgical practice around the world, largely through the dissemination of the meticulously written and carefully illustrated six-volume *Medical and Surgical History of the War of the Rebellion*. (Barnes, 1870)

Both during and after the war, opinion about the effectiveness and quality of civil war surgeons varied widely. One finds references to the dedication and professionalism of surgeons, and also references to surgeons as "butchers" and "quacks." Debate continues to this day about the effectiveness of Civil War surgery, and while several historians[1] present a favorable impression of civil war surgeons, there remains little direct statistical evidence supporting the view that Civil War surgery was effective in improving wounded soldiers' chances of survival. Indeed, popular culture continues to associate Civil War surgery with a degree of horror and crudity, as captured in scenes in popular films such as *Gone With the Wind*[2] and *Dances With Wolves*.[3]

Ideally, one would assess the efficacy of Civil War surgery using observational data that allows comparison of recovery rates in a sample of wounded soldiers,

[1] See Rutkow (2005), Bollett (2002), and Freemon (2001), for example.
[2] https://www.youtube.com/watch?v=-rfa_6pCJ4c
[3] http://www.tcm.com/mediaroom/video/317385/Dances-With-Wolves-Movie-Clip-Coffee-Up.html

within which some wounded received surgery and some did not. Unfortunately, most of the voluminous data gathered during the war on wounds, surgery, and recovery are not set up in such a way so that they might be used to directly assess the effectiveness of treatment.

One exception is a data set assembled by Dr. Edmund Andrews, which to my knowledge has not been analyzed using modern statistical methods. Dr. Andrews served as the Surgeon for the First Regiment of Illinois Light Artillery, and was also Professor of Surgery at Lind University.[4] With the aid of his colleagues in the surgical detachment of the 3rd Division, 13th Core, Army of the Mississippi, under General William T. Sherman's command, he created a unique and interesting document: 'A Complete Record of the Surgery of the Battles Fought Near Vicksburg, Dec. 27, 28, 29, and 30, 1862.' (Andrews, 1863) Dr. Andrews stated his reasons for constructing this document:

> It was with intense chagrin that I...saw the entire loss of scientific results from the bloody battles of Fort Donelson, Shiloh, and the numerous lesser combats in front of Corinth. It is a painful fact, that after these battles the results of the various operations and injuries remained entirely unknown to the original operators, and they gained almost nothing by their experience, except the skill of hand acquired in their manipulations.
>
> For this reason, I resolved at the next large battle...to make a determined effort to secure the entire surgical history of the wounded up to the latest period which the circumstances would permit. In this endeavor I have been successful. (Andrews, 1863, p. 4)

Dr. Andrews and his colleagues did indeed secure, or at least well-approximate, "the entire surgical history of the wounded" for the next large battle: the Battle of Chickasaw Bayou. They recorded and described injuries, the nature of treatment

[4]Dr Andrews had an illustrious career after the war. Lind University later became Northwestern University Medical School, of which Dr. Andrews was a founding member. Dr. Andrews was a pioneer in medical data collection, in the development of medical instruments, and in the use of anesthesia during surgery. He was also an expert geologist, publishing a popular text *The Early Glacial History of North America*.(Northwestern Medical Magazine, 2015)

(if any), and the condition of the patient 15 to 20 days after injury, because by that time, as Andrews put it, “the question of life or death is usually settled.” (Andrews, 1863, p. 32) While the data set is small (n=499), with a little work the data do provide some means for controlling for wound severity, and also for selection into surgery.

My analysis of this data suggests that surgeons increased the typical wounded soldier’s probability of survival. I estimate the Average Treatment Effect (ATE) to be .25-.28. If one focuses on those who received surgery (Treatment on Treated), the estimate increases to .31-.45. In many cases this amounted to a doubling of the odds of survival, so that a wounded soldier with a probability of survival of .3 without surgery would on average have a probability of .6 of survival with it.

Before describing these results and how I arrive at them, it is useful to discuss surgical methods during the war and the continuing debate about the effectiveness of Civil War surgeons.

2 Surgery in the Civil War

The methods for treating wounds during the Civil War, at least at the war’s beginning, derived in large part from the experience of continental surgeons in contemporaneous conflicts such as the Crimean War. Techniques evolved rapidly in response to the unique challenges presented by Civil War wounds.[5] Gunshot wounds were the most common type of wound treated, comprising some 94% of wounds (Wilbur, 1998, p. 45). The cone-shaped Minié ball was the primary ammunition of the Civil War. Minié balls deformed or flattened on impact, and then had a tendency to tumble as they passed through flesh, “tearing a terrible swath through muscle and bone.”(Wilbur, 1998, p. 45). In terms of what this implied for medical procedure; according to one Union surgeon “The minie [sic] ball striking a bone does not permit much debate about amputation,”(Dimon, 1960, cited in Wilbur (1998) p. 45). While amputation figured prominently among available procedures, surgeons also

[5]There are a number of recent books which vividly describe surgery and medicine in the civil war. See, for example, Bollett (2002), Rutkow (2005), and Wilbur (1998) for general discussions of Civil War medicine. Freemon (2001) focuses specifically on surgery, while Wilbur (1998) describes surgical instruments and procedure in great detail.

cleaned and debrided wounds, removed damaged tissue, tied arteries, and if at all possible removed bullets from wounds. (Bollett, 2002, p. 91-2) While surgeons honed methods for dealing with wounds over the course of the war, surgeons did not develop a precise understanding of the primary threat to recovery after surgery: post-operative infection.[6] Doctors did realize, however, that primary amputations (amputations done within a day or so of the injury) reduced the chances of infection, as opposed to more invasive procedures such as resection (removing a damaged section of the bone but otherwise leaving the limb intact).

2.1 Opinions about surgery

The debate about the effectiveness of Civil War surgical care occurred in many arenas with participants from across the social spectrum. Soldiers, nurses, and medical professionals chimed in with opinions about care, as did the editorial pages of many newspapers.

Even though the nature and delivery of care improved over the course of the war, a few factors seem to have instilled a poor popular impression of wartime surgery. One factor was the highly-publicized medical disasters which accompanied Union military disasters in the first year of the war. The Union rout at the 1st Battle of Bull Run (Monasses) in late July, 1861 quickly evolved into a medical nightmare. There was no plan in place to treat or evacuate the wounded prior to the battle, and the medical corps were quickly overwhelmed by the large number of wounded. Wounded soldiers walked miles to receive care if possible - sometimes all the way to Washington - while others, unable to walk, remained where they fell, sometimes for days.(Wilbur, 1998, p.35-7) Even if a soldier managed to get himself to a doctor, the doctor was often ill-prepared in training and equipment to treat him. A New York Times editorial from July 6, 1862 openly agitated for reform of the medical staff, and indicated impressions of care had still not changed almost a year into the war and after several large engagements:

> It is notorious that our brave fellows have...suffered severely for want

[6] Bollett (2002, p. 197) gives 92% as the fatality rate from "Pyæmia," which was the term used to describe infections traveling through the bloodstream.

> of medical officers enough...and from the incompetency of the surgeons...Quacks of all sorts, apothecary's boys, and even barbers and others, wholly destitute of professional knowledge actually hold the posts of Surgeons...("Reform of the Medical Staff", 1862 July 6, *New York Times*)

The situation started to improve with the arrival of the dynamic, energetic Jonathan Letterman with McClelland's Army of the Potomac in July 1862.[7] Even so, some still took exception to the performance and abilities of surgeons. A New York Times Editorial from October 19, 1862 lamented the fact that so many "quacks and butchers", who would perform "hazardous operations...with as little hesitancy as they would carve a joint of beef," had answered a summons for additional Surgeons. The editorial did, however, take pains to differentiate these volunteers from the "eminent surgeons from New York and Philadelphia", whose "professional skill was only equaled by their patriotism and courage."

While organization and treatment became more efficient as the war continued, undergoing or even witnessing surgery was, to put it mildly, an unpleasant experience. This made an impression on the soldiers, and the nature of the task may have colored impressions of those who performed it, whether or not surgery was necessary or done well. As Freemon writes: "Most soldiers of both armies had extensive experience with the butchery of farm animals...[they] observed the surgery performed on others, or upon themselves, and could not help but note the similarity: 'It was butchery, sheer butchery, pure butchery' is the comment of a hundred diaries." (Freemon, 2001, p. 48). To give one example, a private in the 116th Illinois wrote that "The Damed [sic] surgeons are not worth a Curse. They dont no [sic] anything." (Bear, 1961, p. 17, cited in Robertson (1988) p. 158)

As the war pressed on, opinions about surgeons and their actions took a different turn. Surgeons were depicted as dedicated, hard-working, and skilled. Walt Whitman, for one, attested to the "zeal, manliness, and professional spirit and capacity,

[7] Rutkow (2005) points to the battle of Antietam in mid-September, 1862 as a turning point. By this time, some of Letterman's notable innovations in delivery of care, such as his ambulance system, were in place. The innovations in battlefield medicine introduced by Letterman were, in spite of the horrors of the battle of Antietam itself, hailed as "an unqualified success."(Rutkow, 2005, p. 145)

generally prevailing among the surgeons... They are full of genius too."(Whitman, 1962, p. 9-10, cited in Bollett (2002) p. 439) Indeed, Bollett (2002), Freemon (2001), and Rutkow (2005) all present favorable assessments of surgical care during the Civil War from the perspective of the present. Bollett (2002, p. 136) further notes that prominent physicians from Europe were impressed with both the quality and delivery of surgical care by American surgeons during the Civil War.

While the anecdotal evidence is mixed, it is fair to say that many knowledgeable observers had a favorable impression of Civil War surgeons. It is also probably true that, the lack of preparation early in the war and the nature of surgery did much to sour public opinion of surgeons.

2.2 Aggregate Data

The *Surgical and Medical History* contains a large amount of data both on the incidence of surgery and on recovery rates. In Table 1 I present information gleaned from various tables of the *Surgical and Medical History*. This aggregated table admits comparison of the recovery rates from surgery (amputation or excision) and conservation (no aggressive treatment). I constructed the table by combining all tables with this information in the *Surgical and Medical History* with more than 1000 total observations.

The data in table 1 suggest that across a variety of wound types and locations, recovery rates were generally lower among those wounded treated with surgery. This data is, however, limited in its usefulness for drawing conclusions about surgical outcomes because of the almost certain presence of unobserved heterogeneity in the data. Surgeons were likely to intervene in more serious cases, and were inclined to leave lighter, less serious wounds alone. This process of selection into treatment can profoundly impact comparisons.(See, for example Stukel et al., 2007) Unfortunately, there is no way to overcome this in an analysis based on aggregate data like that summarized in table 1. This data does not allow one to control for severity or model the surgical decision.

Other statistical evidence also does not admit definitive conclusions about the effectiveness of Civil War surgery. Bollett (2002, p. 186-8) notes that survival rates

	Cases	Surgery Rate	Recovery Rates		
			Surgery ?		
			No	Yes	Diff.
Hip and Femur Fractures					
Neck of femur	108	11.11	27.08	0.00	27.08
Fracture in upper third of femur	1441	14.09	48.56	27.09	21.47
Fracture in middle third of femur	1256	33.36	58.48	42.00	16.47
Fracture in lower third of femur	1726	66.92	60.48	52.90	7.58
Condyles of femur	593	73.36	43.40	41.84	1.56
Fracture point unspecified	1964	68.64	19.65	34.42	-14.77
Knee Joint					
Patella, not opening joint primarily	145	21.38	71.79	35.48	36.31
Head of tibia	439	69.02	40.88	41.25	-.38
Patella, opening joint primarily	135	65.19	46.81	63.64	-16.83
Part not specified	1797	80.58	23.66	52.42	-28.76
Leg Bone Shot Fractures					
Tibia	2550	24.16	88.08	58.60	29.48
Fibula	1024	22.17	89.45	63.00	26.46
Unspecified	3699	85.97	69.08	57.11	11.98
Tibia and fibula	1424	74.72	78.88	67.39	11.49
Shot Contusions, Fractures: Upper Extremities					
Leg, Ankle	194	11.34	91.28	36.36	54.92
Femur, Knee	205	9.27	75.27	21.05	54.22
Shoulder joint	1440	71.94	70.79	54.54	16.26
Elbow	2743	68.47	88.27	73.64	14.63
Clavier, Scapula	2250	3.73	85.43	71.43	14.00
Wrist	1498	52.94	91.20	82.22	8.98
Skull	328	4.88	83.65	75.00	8.65
Arm	7888	67.00	80.57	73.11	7.46
Hand	9960	83.10	89.94	82.92	7.03
Forearm	5118	43.43	92.22	85.29	6.93
Shot Contusions, Fractures: Lower Extremities					
Leg	8893	56.84	83.62	60.32	23.31
Hip joint	386	21.24	18.09	4.88	13.21
Ankle	1707	70.12	78.86	69.59	9.27
Thigh	6549	47.61	49.72	42.17	7.55
Foot	5419	41.93	87.45	80.59	6.86
Knee Joint	3385	73.77	39.96	48.30	-8.34
Flesh and periarticular wounds					
	59139	0.34	95.27	31.34	63.93

Table 1: Selected information adopted from *Medical History of the Civil War*. The **Surgery Rate** is the rate at which Amputations and Excisions were performed relative to Conservation.

varied as the war went on, and at least over the first three years of the war, fatality rates following surgery fell (although the fatality rate did rise in the final year of the war). Bollett (2002, p. 192) also notes that American soldiers' recovery rates were similar to that of British, French, and German military surgeons at the time, and markedly better than that of Russian Surgeons in the Crimean war. But this data does not say anything about how surgery impacted a wounded soldier's chances of survival.

3 Dr. Andrews's Data

Dr. Andrews's determination to form a full medical record of the next major engagement coincided with the Battle of Chickasaw Bluffs, also called the Battle of Chickasaw Bayou or the Battle of Walnut Hills. The battle took place north of Vicksburg, Mississippi, to the east of the Yazoo River.[8] Roughly a week before the battle, Gen. William Sherman landed north of Vicksburg with an expeditionary force that had traveled down the Mississippi River by boat, with the intention of applying direct pressure on Vicksburg. The ensuing battle and assault on Chickasaw Bluffs opened the Vicksburg campaign.

The battle stretched over four days, and unfolded in two phases. On the first two days of the battle, Union forces pushed Confederate forces eastwards, towards and eventually across Chickasaw Bayou. This phase of the battle was characterized by Union advance and organized Confederate retreat across swamp, bayou, and forest.

After the first two days of action, the Confederates had retreated across Chickasaw Bayou, and had taken up a solid defensive position on the Bayou's opposite side both in front of and on top of the steep bluffs on the eastern side of the Bayou. An assault on this position would mean crossing very difficult terrain composed of swamp, bayou and bluff, and negotiating abatis, rifle pits, and other defensive works. This is how things stood on the evening of December 28th, 1862.

General Sherman ordered a direct assault for the morning of December 29th. Some assaulting forces were to cross existing bridges over the bayou, some were

[8] For the battle details, there are good secondary sources such as Gildner (1991), Jones (2015), and Winschel (2009). There are also good primary sources; see, for example Morgan (1888).

to cross on constructed pontoon bridges, and others were to attempt to ford the Bayou and its contiguous source to the south, McNutt Lake. The assault began behind schedule, and quickly bogged down. While lack of coordination and other logistical difficulties among the assaulting regiments played a role, so did the difficult circumstances and solid defensive positions of the Confederate troops.[9] Many of the assaulting troops foundered in the swampy terrain, while others found the Bayou too deep to wade across. Others managed some success, only to encounter withering fire from dug-in Confederate positions.

On the Union side, casualties were as follows: 212 killed, 1004 wounded, and 500 missing or taken prisoner. Confederate sources reported a total of 213 casualties on their side.

4 Data

Andrews's data set contains information on a total of 730 wounded soldiers. Out of this total, Andrews and his colleagues recorded a 15-20 day-later outcome for 499 observations. One point of interest is that the total of 730 cases is significantly less than the 1004 reported wounded at the battle in the official records. This discrepancy is likely because casualty reports were officers' responsibility and not that of the division medical staff. Therefore, the 1004 wounded may have included some cases that never appeared at the battlefield hospital.

The reason why only 499 cases of the 730 have outcomes, apart from having a listed "unknown" outcome, is that many cases were judged at the division hospital to be lightly wounded enough to return to their regiments. For example, after listing all cases of "Wounds of the Fore-arm" Andrews writes that "with the above should be reckoned 8 cases of slight wounds of forearm...which remained with their regiments." He does not record outcomes for these lightly-wounded cases, but one imagines that in these cases the outcome was almost always favorable. A total of 152 such cases were present in the data, which, when subtracted from 730, leaves

[9] During the night, some troops were to build a pontoon bridge across McNutt Lake but wound up placing the bridges across a different body of water. By the time the error was discovered, it was too late. Efforts to build the bridges in the correct location in the early morning hours of the 29th were deterred by perceptive Confederate troops.

a total of 578 hospital cases. Of these, 79 cases had either no 15-20 day outcome reported, or reported an outcome of “unknown,” resulting in a data set of 499 non-missing observations.

For the 578 hospital cases, Dr. Andrews and his colleagues recorded soldiers’ initials, regiment, injury location, a brief description of the injury, the treatment (if any), and an assessment of the patient’s condition after in most cases 15-20 days. For example, case No. 25 listed under “Wounds of Fore-arm” is for a soldier with initials “R. A. S.” from the 13th Illinois. The soldier’s wound is described as a “compound fracture of the middle forearm,” which was treated with a primary amputation. The soldier was reported to be “doing well, 20th day.”[10](Andrews, 1863, p. 19)

4.1 Outcomes and operations

Based on this data, what fraction of the wounded at Chickasaw Bayou received surgery, and what fraction eventually recovered? If an amputation, resection, or bullet removal is mentioned in the case remarks, I record this as a surgical case (surgery=1); where otherwise I assume surgery did not occur (surgery=0). I also construct a dichotomous outcome variable. If the case was reported as “doing well” or “doing tolerably well” or even “getting better,” I recorded this as indicating the patient survived (survived=1). By contrast, if the wounded patient was reported to have died, or was “not doing well” or “tending towards gangrene” I recorded that the patient did not survive (survived=0).

Table 2 shows cross-tabulations of surgery and outcomes. The top panel of table 2 shows cases with known outcomes in the hospital data set. The overall recovery rates are similar across those who did and those who did not receive surgery, with a slight tendency towards a higher recovery rate among those receiving surgery. This initial result contrasts with the aggregate numbers reported in table 1, in which recovery rates seem to be substantially lower in surgical cases. One might hazard a guess as to why this is so - that non-surgical cases were on average less serious

[10]Andrews also recorded whether or not anesthetic was used if a procedure was performed. As it happens, anesthetic - either Ether or Chloroform - was used on virtually all surgical cases in the data.

outcome	(1) All cases Freq (Percent)	(2) No Surgery Freq (Percent)	(3) Surgery Freq (Percent)
Did not survive	109	82	27
	(21.84)	(23.56)	(17.88)
Survived	390	266	124
	(78.16)	(76.44)	(82.12)
Total	499	348	151

Hospital data only

outcome	(1) All cases Freq (Percent)	(2) No surgery Freq (Percent)	(3) Surgery Freq (Percent)
Did not survive	109	82	27
	(17.44)	(17.26)	(18)
Survived	516	393	123
	(82.56)	(82.74)	(82)
Total	625	475	150

All data, including cases sent back to regiment

Table 2: Data on operations and outcomes from Andrews's Data. The top panel includes only hospital cases. The bottom panel adds cases returned to the regiments without treatment as part of the non-operated, recovered group.

wounds. This might be partially borne out by the lower panel on figure 2. In this panel, I include the slightly wounded cases that were returned to their regiments as non-surgical cases with favorable outcomes. This inclusion renders the recovery rates identical across the surgical and non-surgical groups. The advantage would probably tip further in favor of non-surgical cases if cases that never appeared at the hospital were also included, and official casualty statistics (documented below) suggest that a non-negligible fraction of wounded did not go to the hospital. This reveals an advantage of Andrews's data relative to the aggregate data: his sample renders the sampled cases somewhat closer in terms of their unobserved characteristics by omitting the most lightly wounded who would almost never receive surgery, and who probably never went to the hospital.

4.2 Wound location and severity

Dr. Andrews also recorded where the wounds occurred: either neck-trunk-shoulder, arm, forearm, hand, thigh, knee, leg, or foot. I classify these wounds based on location into how far away they are from the core, using a "distance from core" measure. On the scale, wounds to the neck, trunk, and shoulder and wounds receive a value of zero, while wounds of the hands and feet receive a score of 4. Intermediate wounds to the thigh (score of 1), or knee (score of 2) get a score based on how far they are from the core. I then normalize the resulting distance scale to run on the unit interval. Using a scale variable economizes on the number of variables introduced into the analysis. Introducing dummy variables for wound locations, for example, quickly consumes degrees of freedom that are valuable commodities in a small data set. It also bears mentioning that I did not decide on the distance-from-core scale blindly, but adopted it because it jibes well with Andrews's description of the surgical decision, which I discuss in the next section.

The case descriptions also allow me to form a wound severity measure. I used a count of words in the wound description that indicate the wound was severe. These words included "bad" or "badly", "shattered," "splintered," "compound," "torn," with the word "severe" itself getting a +2 on the scale. I subtracted points from the scale that indicated the wound was not severe. These words included "slight," "flesh wound," and even "finger." Once all these total were added up, I normalized the resulting severity scale to the unit interval.

4.3 Soldier characteristics

While the data offers only limited information about soldiers, it does report regiments and soldiers' initials. A soldier's regiment is useful to know because some regiments were involved in the assault on December 29th, 1862, and therefore produced a large number of wounded in a narrow window of time. I reason that the number of aggregate wounded in a regiment is a useful instrument for the surgical decision. Surgeons had to allocate their scarce time wisely in these situations, and presumably would focus efforts on cases where they would have the most impact. In this setting, some soldiers may not have received surgical treatment that otherwise would have. I therefore create a variable for the aggregate number of observations from each soldiers' regiment as a measure of a triage effect.

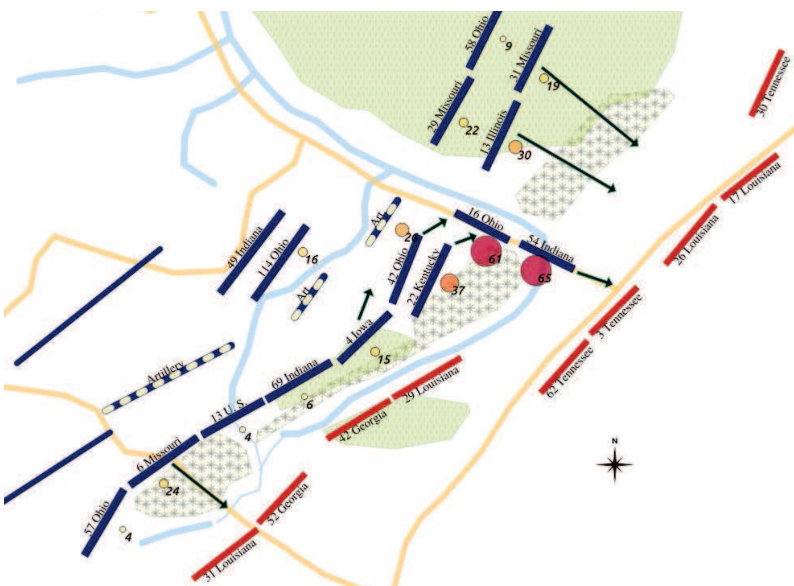

Figure 1: The assault on December 29th, 1862, with regiments marked showing total number of cases in the data. The figure was adopted from Winschel (2009, p. 17) Numbered circles denote the number of observations each regiment contributes to Andrews's data set.

To get a feel for how this instrument plays out in the data, consider figure 1, which shows a map of the assault on December 29th, along with the numbers of observations in the data from each regiment. The 16th Ohio and the 54th Indiana, both of which advanced in double column formation parallel to Chickasaw Bayou, along a road and over a bridge - contribute 61 and 65 cases, respectively, to the data. Other regiments well-represented in the dataset include the 13th Illinois, 31st Missouri, 22nd Kentucky, 42nd Ohio, and 6th Missouri, all of which figured prominently in the assault.

I break down soldiers' initials into a simple dichotomous variable indicating whether or not soldiers had more than two initials.[11] My hope is that the presence of a middle initial, in particular, functions as a proxy for status, and that soldiers of higher status would have been more likely to successfully debate the necessity of surgery with the surgeon. Wilson (2000, p. 300) provides some support for the idea that people of higher status more frequently had middle initials in the middle of the 19th century, claiming that middle names first became popular in the United States at the end of the 18th century among elites then "percolat[ed] downwards," so much so that by the end of the 19th century, virtually everyone had a middle name.

I took a 1% sample from the 1860 U. S. census data provided by IPUMS (Ruggles et al., 2015) to provide some supporting evidence that having a middle initial might correlate with status. The 1860 census provides first (including middle) and last names of respondents, literacy, and industry of employment. Of 221,578 observations in the sample with complete data for all three variables, roughly 21 % had middle names (48,839). The correlation between "employed in a white-collar/professional industry" and "presence of a middle name or initial" is 0.0764 ($p<0.0001$), while the correlation between the census's four-point literacy scale and "middle initial" is .1467 ($p<0.0001$). This offers some support for the idea that those of high social status might have more initials. Bollett (2002, p. 157) notes that "...at the time there was much less acceptance of people with obvious impairments in 'polite society."' One might therefore guess that higher-status soldiers would be more inclined to argue against surgery.[12]

[11] I also attempted to match initials to soldiers' records and names, but a match from just initial was in most cases not possible.

[12] As an instance of a soldier arguing against an amputation, Bollett (2004) documents the inter-

Table 3 presents summary statistics for these variables, and also includes dummy variables for regiments with more than 20 cases in the data. The data in columns 2 and 3 split the data by outcome, while columns 4 and 5 split the data by whether or not surgery was performed.

From table 3 one can see that mean wound severity is higher among both cases that did not survive and cases receiving surgery. Mean distance of the injury from the core is higher among survivors, and among those receiving surgery. The summary data also suggest that surgery were less frequent if a regiment had a larger number of casualties, and a slight tendency for operations to be negatively impacted by the presence of more than two initials. The dummy variables for regiments allow one to see which regiments are most heavily represented in the data. The 54th Indiana and the 16th Ohio spearheaded the assault on December 29, crossing a narrow bridge to engage confederate forces at the base of the bluff. Not surprisingly, these two regiments form the largest fraction of the data, but are followed by other assaulting regiments, including the 22nd Kentucky, the 42nd Ohio, and the 13th Illinois.

5 Models

In the data surgery and survival are both dichotomous variables, so I model the operating decision as based on a latent index s which depends upon wound characteristics X and instruments Z, and an unobservable error term. I observe o^*, which indicates that surgery (an "operation") occurred. Survival s^* depends upon whether or not a latent index s is positive, which itself depends upon wound characteristics X, and o^*; whether the wounded soldier received surgery. A means of modelling the joint incidence of surgery and survival in terms of latent and observed variables is as follows:

esting case of Col. Thomas Reynolds, an Irish immigrant. Col. Reynolds argued (successfully) that his wounded leg should be saved in part because it was "imported."

VARIABLES	(1) All cases mean (sd)	(2) Did nor survive mean (sd)	(3) Survived mean (sd)	(4) No surgery mean (sd)	(5) Surgery mean (sd)
severity	0.518	0.604	0.494	0.465	0.640
	(0.197)	(0.177)	(0.196)	(0.174)	(0.193)
dist. core	0.406	0.284	0.440	0.361	0.510
	(0.365)	(0.329)	(0.368)	(0.362)	(0.351)
¿2 initials	0.309	0.312	0.308	0.313	0.298
	(0.462)	(0.465)	(0.462)	(0.464)	(0.459)
casualties	28.38	28.33	28.40	30.96	22.45
	(22.58)	(22.47)	(22.64)	(22.49)	(21.71)
22nd Kentucky	0.0741	0.0917	0.0692	0.0833	0.0530
	(0.262)	(0.290)	(0.254)	(0.277)	(0.225)
16th Ohio	0.122	0.0826	0.133	0.135	0.0927
	(0.328)	(0.277)	(0.340)	(0.342)	(0.291)
54th Indiana	0.130	0.156	0.123	0.149	0.0861
	(0.337)	(0.364)	(0.329)	(0.357)	(0.281)
42nd Ohio	0.0521	0.0917	0.0410	0.0546	0.0464
	(0.222)	(0.290)	(0.199)	(0.228)	(0.211)
4th Iowa	0.0301	0	0.0385	0.0345	0.0199
	(0.171)	(0)	(0.193)	(0.183)	(0.140)
13th Illinois	0.0601	0.0275	0.0692	0.0632	0.0530
	(0.238)	(0.164)	(0.254)	(0.244)	(0.225)
29th Missouri	0.0441	0.0183	0.0513	0.0575	0.0132
	(0.205)	(0.135)	(0.221)	(0.233)	(0.115)
Survived	0.782			0.764	0.821
	(0.414)			(0.425)	(0.384)
Surgery	0.303	0.248	0.318		
	(0.460)	(0.434)	(0.466)		
Observations	499	109	390	348	151

Table 3: Summary statistics in aggregate, by outcome and by surgery.

$$o^* = \mathbf{1}(o > 0), \quad o \quad = \eta Z + \epsilon_o \tag{1}$$

$$s^* = \mathbf{1}(s > 0), \quad s \quad = \beta X + \delta o^* + \epsilon_s \tag{2}$$

Wound characteristics in X include the measure of wound severity described in the previous section, and also the location of the wound captured in my "distance from core" measure. Dr. Andrews, in his report, indicated that "Wounds of the head, neck, and trunk, from their nature, seldom admit of much surgical assistance."(Andrews, 1863, p. 34) But then Dr. Andrews wrote that "The very opposite is true of the wounds of the extremities," as for these wounds, surgical intervention may "produce vast changes in the ratio between mortality and recovery."(Andrews, 1863, p. 34). Dr. Andrews also notes the necessity of intervention in other cases, such as knee wounds in which there has been joint and/or bone damage.(Andrews, 1863, p. 37)

Instrumental variables Z for the decision to perform surgery include those variables in X, and also the (log of) regimental casualties, and the presence of a middle initial. I assume the marginal distribution of the error terms is standard normal, so that s and o follow a bivariate normal distribution. o^* and s^* can then be modelled using a bivariate probit model as described by Heckman (1978) and Maddala (1983) can be estimated directly by maximum likelihood, and, as Wilde (2000) points out, is identified in some situations in which linear IV estimators are not. One assumes that the error terms in (6) are jointly distributed as:

$$\begin{bmatrix} e_s \\ e_o \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1, \rho \\ \rho, 1 \end{bmatrix} \right) \tag{3}$$

As shown by Greene (2008), one can estimate (6) by maximum likelihood using the bivariate normal distribution, where the log-likelihood contribution of an observation can be written as:

$$LL = \ln \Phi_2(q_s(\beta X + \delta o^*), q_o(\eta Z), q_s q_o \rho) \tag{4}$$

where q_s and q_o in ref (4) are given by:

$$q_s = 2s^* - 1; \quad q_o = 2o^* - 1 \tag{5}$$

One usefulness of the model in (6) and (4) is that it allows correlation across error terms. In the current application, it would not be surprising to find that ρ is negative. This would occur if cases that were more likely to receive treatment were also less likely to have good outcomes, as may be expected if there was something about injury severity that was not picked up in the data. The model can be made more flexible by allowing a full set of interactions between surgery and case characteristics. A typical model of outcomes in this case is a switching regression as advanced originally in Roy (1951) and fleshed out in Heckman and Honore (1990).This model can be written as follows:

$$\begin{aligned} s^* &= \mathbf{1}(s > 0), \qquad s = \begin{cases} X\beta_1 + \epsilon_{1s} & \text{if} \quad o^* > 0 \\ X\beta_0 + \epsilon_{0s} & \text{otherwise} \end{cases} \\ o^* &= \mathbf{1}(o > 0), \qquad o = Z\eta + \epsilon_o \end{aligned} \tag{6}$$

This is effectively a three-variable model with a (latent) outcome for those receiving surgery, an outcome for those not receiving surgery, and a latent variable determining whether surgery occurred. The error structure is:

$$\begin{bmatrix} e_{s1} \\ e_{s0} \\ e_o \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{10} & \rho_1 \\ \rho_{01} & 1 & \rho_0 \\ \rho_1 & \rho_0 & 1 \end{bmatrix} \right) \tag{7}$$

The practical problem estimation of (7) presents is that for each case, only one outcome is observed. There is thus no basis for estimating the parameter ρ_{01}. While there is a cottage industry in developing methods to learn about the parameter ρ_{10} (see, e.g., **?**), a common practical solution to the problem - the one adopted in this paper - is to assume ρ_{10} is zero. In this case, the new model can be estimated in much the same fashion as 6, using a bivariate probit function, since with $\rho_{01} = 0$ reduces things to bivariate probits conditional on surgery or no surgery. To be precise, the model is estimated as:

$$LL = \ln \Phi_2(q_o \left[o^*(\beta_1 X) - (1-o^*)\beta_0 X\right], q_o(\eta Z), q_s \left[o^*\rho_1 - (1-o^*)\rho_0\right]) \tag{8}$$

where q_s and q_o in ref (4) are as given in (5). Practically speaking, the model in (8) is essentially just a bivariate probit model in which interactions with the surgery variable and other explanatory variables are included in the model, and the ρ parameter is allowed to vary with the occurrence of surgery.

There are other alternatives for estimating models with dichotomous outcomes and treatments. These methods are approximate in that they sidestep, or ignore, the dichotomous nature of the outcome and/or the treatment variable. Residual inclusion/control function methods, in which one first estimates a model of the surgery decision, obtains residuals from this model, and includes (a function of) residuals in a second-stage model of survival, violate key model assumptions. These models require prediction errors from a treatment model to be normally distributed, which they cannot be in the case of a dichotomous endogenous variable. Angrist (2001) and Angrist and Pischke (2008) mount a compelling defense of using simple linear instrumental variables (IV) methods, but application of linear IV methods in all cases may not be appropriate or desirable. Nichols (2011), in his description of pitfalls and nuances of estimation of models with binary response and treatment, catalogs reasons for not opting for simpler methods. Relative to IV methods, maximum likelihood presents a large gain in efficiency, which may be critical when the sample size is small. Furthermore, linear IV can only uncover average treatment effects across the sample. If one is concerned about selection into treatment, or heterogeneity in impact, one would prefer an estimation method that allows fuller characterization of the joint distribution of outcomes and treatments. As the sample in this paper is small (N=499), and there is at least the possibility that unobserved heterogeneity influences the surgery decision, I favor the bivariate probit and regression-switching models, but also present results for a linear IV model and an instrumental-variables probit model estimated by residual inclusion .

5.1 The decision to operate

Before investigating simultaneous models of surgery and survival, it is helpful to assess how the surgical decision was influenced by soldier and wound characteristics. Accordingly, I begin by focusing on simple models of the decision to operate, following equation (6). Some simple single-equation probit models of the decision to perform surgery on a wounded soldier are presented in table 4.

	(1)	(2)
	I	II
VARIABLES	Surgery	Surgery
severity	2.802***	2.740***
	(0.334)	(0.348)
dist. core	0.663***	0.746***
	(0.173)	(0.176)
ln(reg. cases)		-0.229***
		(0.0567)
$>$2 initials		-0.166
		(0.141)
Constant	-2.323***	-1.626***
	(0.217)	(0.281)
Observations	500	500

Robust standard errors in parentheses
*** p$<$0.01, ** p$<$0.05, * p$<$0.1

Table 4: Modeling the incidence of surgery with single-equation probit models. The dependent variable =1 if surgery was performed.

The first Probit model in table 4 omits the (log of) regimental cases and the middle-initial dummy variable, while the second includes both variables. The (log of) the number of regimental cases suggests a significant triage effect; a wounded soldier from a regiment with a lot of other wounded soldiers appears to have been less likely to receive surgery. The presence of a middle initial has a negative but imprecise impact on the surgical decision. The partial F-statistic on these two variables is 17.75 deriving from the models in table 4, suggesting that they have significant explanatory power in predicting the incidence of surgery, independent of wound severity and location.

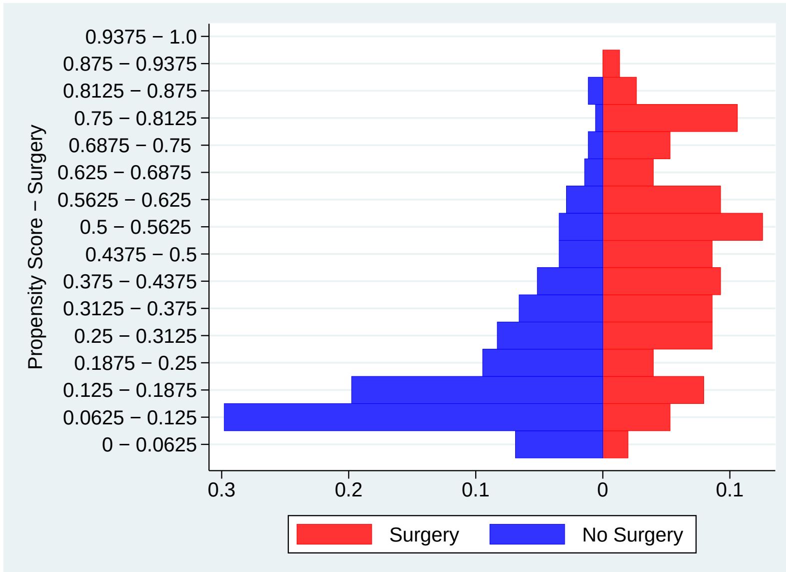

Figure 2: Supports for surgical and non-surgical groups.

Using the generated propensity scores from this model, one can get some sense as to the overlap between the surgical and non-surgical group in terms of the propensity to have surgery. The joint distribution of treatment incidence and propensity scores deriving from the second model in table 4 are plotted in figure 2.

One can see from the figure that there is a difference in the distribution of propensity scores across the two groups. The non-surgical group has a large proportion of cases that were unlikely to get surgery. These are typically soldiers with relatively light wounds. While distributions differ, it is useful to note that there is significant overlap in the distribution for a meaningful range of scores.

5.2 Outcomes

I now describe some naíve models of survival following equation (2), in which surgery is taken to be an exogenous variable. While these models do not take into account model endogeneity, they can lend insight into other aspects of the efficacy of Civil War surgery. One can, for example, see how controlling for wound severity and location figures into an assessment of surgical effectiveness, which is not possible in the aggregate data. Single-equation probit models of survival appear in

VARIABLES	(1) I Survived	(2) II Survived	(3) III Survived
Surgery	0.199		0.490***
	(0.141)		(0.175)
severity		-1.833***	-2.286***
		(0.318)	(0.371)
dist. core		0.822***	0.705***
		(0.203)	(0.203)
Constant	0.720***	1.472***	1.616***
	(0.0740)	(0.184)	(0.198)
Observations	499	499	499

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Single-equation probit models of outcomes showing impact of surgery, both with and without controlling for severity and location.

table 5.

The first model on table 5 is included to show what happens if one estimates a model of surgical impact without controlling for wound characteristics. These results suggest that surgery is ineffective. From this one might intuit that Civil war data on surgery and recovery has the same selection bias problems that many modern observational data sets do: more severe cases tend to select treatment. By contrast, in model II severity and distance from core are included in the probit model, and surgery now appears to have a significant and positive impact on the outcome.

5.3 Joint models of surgery and survival

Simultaneous models of surgery and survival are displayed in table 6. Model I is a bivariate probit model, while Model II is a switching regression model, where a full set of interaction terms is included. A couple of points of interest in table 6 are as follows. The bivariate probit (Model I on table 6) has a larger coefficient on the Surgery variable than the single-equation counterpart (model III on table 5), sug-

		(1)	(2)
		I	II
EQUATION	VARIABLES		
Survived	Surgery	1.135**	-1.573
		(0.528)	(1.459)
	severity*operated		3.294**
			(1.409)
	dist. core*operated		-0.0220
			(0.503)
	severity	-2.848***	-3.442***
		(0.530)	(0.414)
	dist. core	0.521*	0.436
		(0.272)	(0.326)
	Constant	1.751***	1.918***
		(0.214)	(0.228)
Operated	severity	2.815***	2.841***
		(0.338)	(0.340)
	dist. core	0.736***	0.758***
		(0.175)	(0.175)
	ln(reg. cases)	-0.225***	-0.229***
		(0.0581)	(0.0566)
	$>$2 initials	-0.155	-0.133
		(0.142)	(0.136)
	Constant	-1.675***	-1.696***
		(0.270)	(0.271)
Corr.	Surgery		1.163**
			(0.537)
	Constant	-0.423	-0.706***
		(0.356)	(0.261)
	Observations	499	499

Standard errors in parentheses
*** p¡0.01, ** p¡0.05, * p¡0.1

Table 6: Simultaneous models of surgery and survival. Model I: Bivariate Probit. Model II: Switching regression.

gesting that one would underestimate the impact of surgery if the selection process was not accounted for. The estimated cross-equation correlation coefficient ρ is also negative (although imprecisely estimated), suggesting that even after controlling for wound characteristics, a tendency for unobserved characteristics of wounds to lead to both surgery and non-survival.

For purposes of comparison, tab 7 includes the approximate models. While the models suggest outcomes that are quantitatively similar to those in table 6 (taking care to remember that the coefficients in column II of 7 pertain to a probit), the estimated coefficient on Surgery is insignificant. This could be due to the efficiency losses, or could be due to features of the distribution of treatment muting the estimated effect. There is some evidence that the latter effect is important.

In table 8, I describe treatment effects deriving from models I and II on table **??**. These effects are computed for the bivariate probit and switching regression models on table 6.[13] From table 8, one sees that effects seem to be larger among those who received surgery than among those who did not. This suggests there was some selection into surgery based on some unobserved aspect that made treatment of the wound by surgery more likely to have a good outcome. The differences in the predicted efficacy of surgery are especially stark in the predictions from the switching regression model.

Figure 3 plots the difference in survival probability as a function of the probability of treatment for each of the bivariate normal models. From the figure, it seems as though a lot of the cases that did not receive surgery would not have benefitted from it, and in fact may have been harmed by surgery, while the bulk of those who received surgery benefitted from it a great deal. This is especially true for the predictions deriving from the switching regression. This might also suggest a reason for the imprecision of the linear IV estimates; there is a rather substantial difference between how the average wounded soldier would have responded to surgery relative to those who actually received it.

[13]Effects are computed by predicting the probabilities of survival if surgery=1 and if surgery=0, taking the difference for each case, and then averaging across cases.

EQUATION	VARIABLES	(1) I	(2) II
Survived	Surgery	0.207	0.899
		(0.226)	(0.815)
	severity	-0.698***	-2.649***
		(0.223)	(0.750)
	dist. core	0.167**	0.608**
		(0.0693)	(0.290)
	Constant	1.013***	1.707***
		(0.0746)	(0.247)
Surgery	severity	0.876***	0.884***
		(0.0944)	(0.0941)
	dist. core	0.230***	0.226***
		(0.0500)	(0.0498)
	ln(reg. cases)	-0.0660***	-0.0663***
		(0.0167)	(0.0166)
	>2 initials	-0.0429	-0.0424
		(0.0397)	(0.0397)
	Constant	-0.0382	-0.0397
		(0.0783)	(0.0778)
Ancillary parms.	rho		-0.176
			(0.358)
	ln(sigma)		-0.905***
			(0.0317)
	Observations	500	499
	R-squared	0.223	

Standard errors in parentheses
*** p¡0.01, ** p¡0.05, * p¡0.1

Table 7: Approximate models of surgery and survival. Model I: Linear IV model. Model II: Residual inclusion/Control Function

VARIABLES	(1) ATE	(2) TT	(3) TUT
Bivariate Probit	0.253	0.313	0.228
	(0.005)	(0.009)	(0.005)
Regime Switching Model	0.0610	0.201	0.000214
	(0.010)	(0.018)	(0.010)
Observations	499	151	348

Table 8: Summary of estimated average treatment effect (ATE), average treatment effect on treated (TT), and average treatment effect on untreated (TUT), with standard errors.

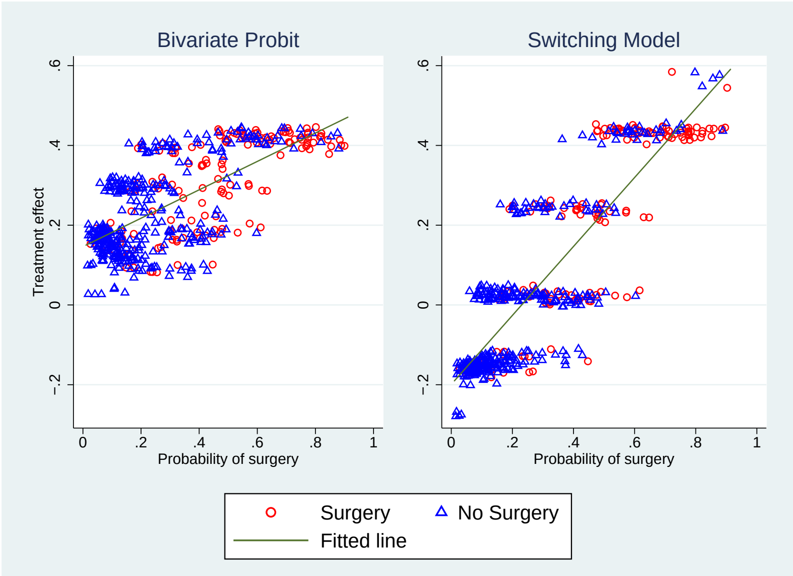

Figure 3: The distribution of predicted outcomes both with and without surgery, as a function of the predicted likelihood of surgery, for both the bivariate probit and the switching regression model.

6 Conclusion

After the war, many prominent surgeons argued the case that civil war surgeries in general, and amputations in particular, saved a great many lives. Some even advanced the view that surgeons were, if anything, too conservative. William Williams Keen, a young Civil War surgeon who went on to become one of the most famous surgeons in the country after the war, wrote "I have no hesitation in saying that far more lives were lost in refusal to amputate than by amputation."(Keen, 1905, p. 443, cited in Bollett (2004)) Jonathan Letterman also opined that "if any fault was committed, it was that the knife was not used enough."(Letterman, 1870, p. 99, cited in Rutkow (2005) p. 174)

The results of this paper suggest that the convictions of Letterman and Keen had some basis. Models that control for severity and account for essential heterogeneity in the data indicate that Civil War surgeons were effective and increased wounded soldiers' chances of survival substantially - on the order of 50-75 % - and perhaps even more for the severely wounded.

In the motion picture *Dances with Wolves*, Lt. John Dunbar (played by Kevin Costner) escapes surgery because the regimental surgeons needed a coffee break just as he arrives on the operating table. While his character eventually recovered from his wound, it is probable that he recovered in spite of his escape from the surgeons, and not because of it.

References

Andrews, Edmund. (1863), *Complete Record of the Battles fought near Vicksburg December 27, 28, 29, & 30, 1862,* George H. Fergus.

Angrist, Joshua D. (2001). 'Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice', *Journal of Business and Economic Statistics* 19, 2–28.

Angrist, Joshua D. and Pischke, Joern-Steffen. (2008), *Mostly Harmless Econometrics: An epiricist's companion,* Princeton University Press, Princeton, New Jersey.

Barnes, Joseph K. (1870), *The Medical and Surgical History of the War of the Rebellion, (1861-1865),* Government Printing Office, Washington, D. C.

Bear, Henry C. (1961), *The Civil War Letters of Henry C. Bear: A Soldier in the 116th Illinois Volunteer Infantry,* Lincoln Memorial University Press, Harrogate, Tennessee.

Bollett, Alfred Jay. (2002), *Civil War Medicine: Challenges and Triumphs,* Galen Press, Ltd., Tucson, Arizona.

Bollett, Arthur Jay. (2004). 'The Truth about Civil War Surgery', *Civil War Times* 43(4), 26–56.

Dimon, Theodore. (1960), A Federal Surgeon at Sharpsburg, *in* **James I. Robertson**., ed., 'Civil War History 6'.

Freemon, Frank R. (2001), *Gangrene and Glory: Medical Care during the American Civil War,* University of Illinois Press, Urbana and Chicago, Illinois.

Gildner, Gray M. (1991), *The Chickasaw Bayou Campaign (Master's Thesis),* U. S. Army and General Staff College, Leavenworth, Kansas.

Greene, William H. (2008), *Econometric Analysis, 6th ed.*

Heckman, James J. (1978). 'Dummy Endogenous variables in a Simultaneous Equation System', *Econometrica* 5, 931–59.

Heckman, James J. and Honore, Bo E. (1990). 'The empirical content of the Roy model', *Econometrica* 58, 1121–49.

Jones, Michael Dan. (2015), *The Battle of Chickasaw Bayou, Mississippi: A Confederate Victory in the Vicksburg Campaign,* Michael Dan Jones.

Keen, William Williams. (1905), *Addresses and Other Papers*, W. B. Saunders, Philadelphia, Pennsylvania and London.

Letterman, Jonathan. (1870), Excerpts from a report of the operations of the medical department of the army of the Potomac from Jully 4th to December 31st, 1862, *in* **Joseph K. Barnes**., ed., 'The Medical and Surgical History of the War of the Rebellion', Vol. 1, Government Printing Office, Washington, D. C., pp. p. 92–104.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

Morgan, Gen. George. (1888), The Assault on Chickasaw Bluffs, *in* **Robert Underwood Johnson and Clarence Clough Buel**., eds, 'Battles and Leaders of the Civil War', Vol. 3.

Nichols, Austin. (2011), *Causal inference for binary regression*, accessed December 2, 2016, http://www.stata.com/meeting/chicago11/materials/chi11_nichols.pdf.

Northwestern Medical Magazine. (2015), *Edmund Andrews, A Medical School Founder*, accessed November 2, 2016, http://magazine.nm.org/2015/8/dr-edmund-andrews-a-mmedical-school-founder.

Robertson, James I. (1988), *Soldiers Blue and Gray*, University of South Carolina Press, Columbia, South Carolina.

Roy, A. D. (1951). 'Some thoughts on the distribution of earnings', *Oxford Economic Papers* 3, 135–46.

Ruggles, Steven, Genadek, Katie, Goeken, Ronald, Grover, Josiah and Sobek, Matthew. (2015), *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]*, University of Minnesota, Minneapolis, Minnesota.

Rutkow, Ira. (2005), *Bleeding Blue and Gray: Civil War Surgery and the Evolution of American Medicine*, Stackpole Books, Mechanicsburg, Pennsylvania.

Stukel, Therese, Fisher, Elliot S., Wennberg, David E., Alter, David A., Gottlieb, Daniel J. and Vermeulen, Marian J. (2007). 'Analysis of Observational Studies in the Presence of Treatment Selection Bias', *Journal of* *the American Medical Association* 297(3), 278–85.

Whitman, Walt. (1962), *Walt Whitman's Memoranda During the War*, Indiana University Press.

Wilbur, C. Keith. (1998), *Civil War Medicine*, The Globe Pequot Press, Guilford, Connecticut.

Wilde, Joachim. (2000). 'Identification of multiple equation probit models with endogenous dummy regressors', *Economics Letters* 69, 309–12.

Wilson, Stephen. (2000), *The Means of Naming: A Social History,* Routledge, London, U. K.

Winschel, Terrence J. (2009). 'Chickasaw Bayou: Sherman's Winter of Despair', *Blue and Gray: For those who still hear the guns* 26(3), 6–52.