# QUEUING FOR SURGERY: IS THE U.S. OR CANADA WORSE OFF?

Barton H. Hamilton, Vivian Ho, and Dana P. Goldman*

*ABSTRACT*—Restricted government spending along with universal health insurance has led to longer queues for surgical procedures in Canada versus the United States. Yet it is unclear whether these treatment delays affect health outcomes. This paper tests this hypothesis by comparing the determinants of wait time for hip-fracture surgery and its impact on postsurgery length of stay and inpatient mortality in Canada and the United States. Hazards for surgery/no surgery and discharge alive versus dead are modeled using a competing-risks model. Day of the week of admission is used to help identify the surgery wait-time distribution. We control for unobserved (to the econometrician) health status which may affect wait times and outcomes by assuming a semiparametric distribution for unobserved heterogeneity. We find that predicted hazards for inpatient mortality are virtually identical in Canada and the United States. Yet wait times for surgery are longer in Canada, and surgery delay has a significant impact on postsurgery length of stay in both countries. However, the magnitude of this effect is small relative to other patient and hospital-specific factors. Focusing attention on treatment delays as a weakness in the Canadian health care system may be misleading policymakers from hospital-specific inefficiencies that may have more-important implications for health care costs and patient welfare.

## I. Introduction

POLICYMAKERS and researchers have long been debating the relative merits of the U.S. and Canadian health care systems. Proponents of the Canadian system of universal health insurance often point to disparities in hospital care by insurance and income in the United States (Hadley, Steinberg, and Feder (1991); Wenneker, Weissman, and Epstein (1990)). Critics counter that the Canadian system has led to unacceptable delays in obtaining health care services in Canada. The existence of such queues has been widely documented (Globerman (1991); U.S. Government Accounting Office (1991); Katz, Mizgala, and Welch (1991)). Still, there is little evidence that these queues affect clinical outcomes. If they do not, then criticisms of Canadian queues may be distracting policymakers from other more-important issues affecting patient outcomes.

In this paper, we compare surgical queues and outcomes between the United States and Canada after a patient has fractured his or her hip. We choose this diagnosis for three reasons. First, hip fractures are common and are, in fact, the leading cause of hospitalization for injuries among the elderly (Baker, O'Neill, and Karpf (1984)). Second, there is a strong a priori presumption that surgical queues for hip fracture will affect outcomes. Unlike queuing for many other surgical procedures, treatment for hip fracture is urgent. Patients who fracture their hips are immobile and must remain in hospital, in traction, and on medication until they receive surgery. Prolonged immobility resulting from delay can lead to complications that are potentially fatal (Sabiston (1991)). For example, bed-ridden patients are prone to fluid buildup in the lungs, which places them at high risk for pneumonia. Thus, hip fractures represent an important case study for comparing queues and clinical outcomes in Canada and the United States. Third, hip fractures require immediate hospitalization, so selection into the queue is not an issue.

We analyze discharge data drawn for 20,995 hip-fracture patients admitted to acute-care hospitals in Quebec and Massachusetts between 1990 and 1992. We examine how wait times for surgery affect two important outcomes: inpatient mortality and postsurgical length of stay. Equity concerns also motivate an examination of whether surgical wait times vary by socioeconomic status. In addition, we distinguish between patient-level and hospital-level effects. Low-income patients may have access only to hospitals that tend to have long delays and poor outcomes. Within the hospital, income may have no effect on delays and outcomes.

Past research has already cast doubt on the hypothesis that patients with longer delays experience worse outcomes in Canada (Hamilton, Hamilton, and Mayo (1996)). However, Canadians may face delays that are longer than those in other countries, so that the detrimental impact of delays on outcomes is not discernible with Canadian data alone. This study advances the literature by explicitly comparing surgery delays and their impact on inpatient mortality and postsurgery length of stay in the United States versus Canada. It is also the first study we know of that compares variability in patient outcomes across hospitals in these two countries. A motivation for universal health insurance is that it promotes equal access to care. However, this hypothesis has not been tested with respect to the care patients receive across different facilities.

We utilize survival analysis to examine the timing of death or discharge as a function of wait time for surgery and other covariates. Wait time for surgery is modeled explicitly, and we rely on the day of the week of admission to help identify the distribution of surgery delays. The analysis is complicated by the fact that postsurgery length of stay and mortality are correlated. In particular, one is less likely to observe an inhospital death for those who leave earlier. We model this correlation using a duration model with multiple destinations. Moreover, observable and unobservable factors are likely to affect both wait time for surgery and postsurgical outcomes. For instance, physicians will sometimes delay surgery to stabilize comorbidities that might otherwise threaten the outcome of the surgery (Kenzora et al. (1984)). Thus, we also assume a flexible distribution for unobserved heterogeneity to account for unobserved (by the

Received for publication February 12, 1998. Revision accepted for publication July 7, 1999.

* Washington University Olin School of Business, Washington University Olin School of Business, and RAND, respectively.

We thank Janet Currie, Stephen Zuckerman, two anonymous referees, and seminar participants at the 1996 winter meetings of the Econometric Society, the Washington University Olin School of Business, the University of Manitoba, and the University of Arizona for helpful comments. This research project was supported by a grant from the Social Sciences and Humanities Research Council of Canada.

*The Review of Economics and Statistics,* May 2000, 82(2): 297–308

researcher) differences across patients at the time of admission.

We find that the predicted hazards for inpatient mortality are virtually identical in Canada and the United States. Nevertheless, wait times for surgery are longer in Canada. Surgery delay also has a significant detrimental impact on postsurgery inhospital stays in both countries. However, simulations reveal that the magnitude of this effect is negligible. We do find much longer postsurgery lengths of stay in Canada, and these vary widely across hospitals. It also appears that low-income Canadians and Americans have access only to hospitals with longer wait times. We conclude from this analysis that focusing attention on treatment delays may be misleading policymakers from hospital-specific inefficiencies that may have more-important implications for health care costs and patient welfare.

Section II describes the data and presents descriptive statistics. Section III outlines the framework for examining the determinants of delay and postsurgery length of stay and inhospital mortality in Canada and the United States. We present our estimation results in section IV, simulations in section V, and concluding remarks in section VI.

TABLE 1.—RELATIONSHIP BETWEEN WAIT TIME UNTIL SURGERY, CASE MIX, AND OUTCOMES FOR PATIENTS UNDERGOING SURGERY, BY COUNTRY

| Surgery Delay (days) | Frac. | Charlson Index | Fraction Died in Hospital | Postsurgery Mean | Length of Stay Median |
|---|---|---|---|---|---|
| Panel A: Canada (Quebec) | | | | | |
| 1 | 0.24 | 0.45 | 0.066 | 25.7 | 15 |
| 2 | 0.42 | 0.52 | 0.073 | 26.8 | 16 |
| 3 | 0.17 | 0.56 | 0.086 | 29.0 | 18 |
| 4 | 0.07 | 0.56 | 0.080 | 28.1 | 18 |
| 5 | 0.03 | 0.68 | 0.120 | 34.3 | 20 |
| 6+ | 0.07 | 0.76 | 0.116 | 38.5 | 22 |
| Total | — | 0.53 | 0.078 | 28.0 | 17 |
| Panel B: United States (Massachusetts) | | | | | |
| 1 | 0.18 | 0.53 | 0.026 | 12.9 | 10 |
| 2 | 0.52 | 0.59 | 0.037 | 13.2 | 10 |
| 3 | 0.16 | 0.65 | 0.062 | 15.3 | 10 |
| 4 | 0.05 | 0.75 | 0.062 | 14.4 | 11 |
| 5 | 0.03 | 0.86 | 0.068 | 13.3 | 11 |
| 6+ | 0.06 | 0.71 | 0.072 | 20.1 | 12 |
| Total | — | 0.61 | 0.043 | 13.9 | 10 |

## II. Data and Preliminary Evidence

The data consist of information for 12,016 hip-fracture patients discharged from all 68 hospitals in Quebec and 8,979 patients discharged from 45 hospitals in Massachusetts.[1] For each patient, we know whether hip surgery was performed,[2] and if so, the number of days between admission and date of surgery. We also know postsurgery length of stay, status at discharge (alive or dead), age, sex, type of hip fracture (transcervical, pertrochanteric, or unspecified),[3] the hospital the patient was admitted to, year of admission, income in postal code of residence, number and type of comorbidities at the time of admission, and the day of the week of admission. In any health-outcomes study, failure to adequately control for illness severity can lead to incorrect inferences. Fortunately, our data contain detailed information on disease comorbidities, which we used to construct a Charlson comorbidity index for each patient (Charlson et al. (1987)). This index has been validated as a predictor of mortality in longitudinal studies.[4] Of the patients in the original Quebec and Massachusetts samples, 2.7% and 3.8% had missing income data and were excluded from the sample.

Table 1 presents some provocative descriptive statistics by country for those patients who received surgery. Of those patients operated on within one day of admission, 6.6% died in Quebec as compared with 2.6% in Massachusetts. As the wait time for surgery increases, the mortality rates also tend to rise, reaching 11.6% and 7.2%, respectively, for patients waiting six or more days. The final two columns of table 1 demonstrate that postsurgery length of stay increases as wait time increases in both countries. Thus, the higher inpatient mortality rates within each delay category in Quebec may be due to the higher probability of observing an inpatient death for any given patient in Canada. (Patients are in the hospital longer, so they are more likely to be observed to die there.) Our empirical work controls for this differential censoring across countries by examining the possibility of inhospital mortality conditional upon length of stay.

Other studies have also found that longer delays are associated with longer postsurgery hospital stays and/or higher mortality (Bredahl et al. (1992); Davis et al. (1987); Hoerer, Volpin, and Stein (1993); Roos et al. (1996)). This type of relationship may have led policymakers and researchers to conclude that delays cost money and lives. However, as the third column of table 1 shows, patients waiting longer for surgery tend to be sicker at the time of admission, based on our comorbidity index. Individuals with more comorbidities may be delayed for surgery until their other medical conditions are stabilized; yet the presence of multiple diseases will also increase their probability of death and potentially increase postsurgery length of stay. Even after controlling for these comorbidities, the relationship between delay and outcomes may still reflect underlying patient

[1] The data set consists of all hospitals in Quebec performing hip-fracture surgery, and a random sample of hospitals in Massachusetts. More details are available in appendix A.

[2] Of the Quebec patients, 92% underwent surgery; in Massachusetts, 91% received surgery. Appendix A provides more details on the data.

[3] Transcervical fractures are most commonly treated with total hip replacement, while pertrochanteric fractures tend to result in pinning of the existing hip. Because the existing hip may still be relatively frail, patients with pertrochanteric fractures may take longer to recover from surgery and become mobile.

[4] Higher values for the Charlson index imply worse health or more-severe illness. The value of the index for each patient was constructed using a coding methodology developed specifically for administrative data (Romano, Roos, and Jollis (1993)).

frailty at the time of hospital admission, rather than any detrimental impact of delays. Our empirical work attempts to control for this type of unobserved heterogeneity.

In addition, the Charlson comorbidity index is higher for U.S. versus Canadian patients as a whole, and for each delay category. The higher number of comorbidities coded in the U.S. data is likely due to DRG creep (classifying patients in the highest possible Diagnosis Related Group); U.S. hospitals have an incentive to report more comorbidities per patient, which will increase their payments under the Medicare Diagnosis Related Group reimbursement system.[5] Thus, we will account for potential differences in reporting of patient health status in Quebec and Massachusetts.

## III. Econometric Framework

This section presents our econometric framework for examining the impact of wait time on postsurgery hospital stays and inpatient mortality. We will estimate the model separately for Quebec and Massachusetts, and then compare the estimates across countries. Because postsurgery length of stay and discharge destination are likely to be correlated, these outcomes are estimated jointly using a duration model with multiple destinations (Lancaster (1990), Kalbfleisch and Prentice (1980)).

We first denote the duration of a postsurgery hospital stay by $t$ and suppose that there exist two mutually exclusive and exhaustive discharge outcomes: discharge alive (a) versus discharge dead (d)—indexed by $r$. Let $\delta_r = 1$ if the patient is discharged to destination $r$, and 0 otherwise. The building blocks of the analysis are the transition intensities:

$$\lambda_r(t) = \lim_{\Delta t \to 0^+} \frac{Pr[t < T \geq t + \Delta t,\ \delta_r = 1 | T \geq t]}{\Delta t}, \quad (1)$$

where $\lambda_r(t)$ represents the instantaneous probability that an individual will depart to state $r$ at time $t$, given survival to time $t$. A convenient way of conceptualizing a duration model with multiple destinations is to assume that there exist latent survival times ($T_a$ and $T_d$) corresponding to each destination. The observed duration of a postsurgery hospital stay ($t$) is then equal to min $\{T_a, T_d\}$, and the observed destination state is argmin $\{T_a, T_d\}$. For example, a patient who dies in the hospital has $t = T_d$, and, therefore, $T_a > t$. In this example, live discharges may be treated as right-censored observations of their associated latent survival times until inpatient death. Allowing the transition intensities to depend upon individual characteristics $X$, which includes presurgery wait time, the probability of observing an exit to $r$ after a hospital stay of length $t$ is

$$f_r(t|X) = \lambda_r(t|X) \prod_{j \in D} \exp\left(- \int_0^t \lambda_j(u|X)\, du\right), \quad (2)$$

where $D$ is the set of destinations. The first term of equation (2) is the transition intensity defined earlier. The second term is the survivor function, which is the probability that the individual survives at least to time $t$ in the hospital and hence did not exit to any destination prior to $t$.[6] Equation (2) gives each patient's contribution to the likelihood function.

### A. *Other Econometric Issues*

The econometric analysis must address three issues. First, wait time, which is subsumed in the vector $X$, may be endogenous (in the sense that there may be a common unobservable component influencing both delay and postsurgery length of stay and/or mortality). Second, not all patients receive surgery.[7] Thus, if there are unobserved factors affecting the decision to perform surgery, then delay times may be subject to sample-selection bias. Finally, we wish to identify the extent to which differences both within and between countries in wait time and outcomes reflect hospital-level versus patient-level variation.

Turning to the first issue, table 1 showed that patients with more preexisting medical conditions had longer delays as well as longer postsurgery lengths of stay and higher probabilities of death. While we can control for differences across patients in the number of preexisting conditions, further unmeasurable differences in health status across patients (such as, frailty, which cannot be fully documented in discharge abstracts) may still affect both wait times and postsurgical outcomes. Failure to account for these differences would lead to biased estimates. Thus, we expand our framework to jointly estimate the probability of observing a wait time of $t_w$ days, a postsurgery length of stay of $t$ days, and a discharge to destination $r$. We follow the common approach of decomposing the unobservables affecting $t_w$, $t$, and $r$ into a component that is correlated across $t_w$, $t$, and $r$, and independent components that are specific to $t_w$, $t$, and $r$ (Lancaster (1990)). We denote the correlated component by the scalar random variable $\upsilon$. In our case, the correlated component is best interpreted as unobserved health status of the patient at the time of admission to the hospital. This heterogeneity will affect both the delay until surgery as well as postsurgical outcomes. We will model the distribution of $\upsilon$ semiparametrically, assuming $\upsilon$ can take any of $K$ discrete values.

We are also concerned about selection as to who undergoes surgery. Approximately 8% of Canadian and 9% of U.S. hip-fracture patients admitted to the hospital in our sample never have surgery.[8] We assume that, at the time of

[5] No such incentive exists in Canada, where hospitals receive an annual global budget that is not directly tied to patient case mix.

[6] We assume in equation (2) that the latent survival times $T_a$ and $T_d$ are independent conditional upon $X$. This assumption is relaxed in the subsequent econometric model.

[7] The majority of these patients were discharged alive without surgery, most likely because they were too frail to undergo the general anaesthesia required for general surgery.

[8] Of these individuals, approximately 22% of Canadian patients and 19% of U.S. patients died in hospital. The remaining patients were discharged alive without surgery, most likely because they were too frail to undergo the general anaesthesia required for hip surgery.

admission, all hip-fracture patients have a positive probability of undergoing surgery before discharge. Consequently, if patients not undergoing surgery systematically differ from those who do, then excluding these patients from the analysis may bias estimates of wait times.[9] A duration framework with multiple destinations analogous to equation (1) and (2) can be used to account for selection bias by jointly estimating delay time and whether the patient undergoes surgery (denoted as destination $s$) or exits the queue without receiving surgery (denoted as destination $n$).

We assume that the probability of observing a delay of $t_w$ days prior to receiving surgery (that is, an exit to destination $s$) depends on a vector of measured characteristics $X_w$ (which includes some variables not contained in $X$) and the unobserved (to the researcher) component $\upsilon$, which also affects postsurgical outcomes. Then the density function for surgery is similar to that for discharge as described in equation (2):

$$f_s(t_w|X_w, \upsilon) = \lambda_s(t_w|X_w, \upsilon)\exp\left(-\int_0^{t_w}\lambda_s(u|X_w, \upsilon)\,du\right) \times \exp\left(-\int_0^{t_w}\lambda_n(u|X_w, \upsilon)\,du\right). \tag{3}$$

The first term on the right side represents the conditional probability that an individual exits the queue to receive surgery in day $t_w$, having been in the queue up to that point, while the latter two terms represent the probability that the patient did not leave the queue for any reason prior to day $t_w$. A similar expression may be constructed for the probability that an individual experiences a stay of $t_w$ days and exits without receiving surgery, $f_n(t_w|X_w, \upsilon)$.

We are now able to construct the likelihood of observing a wait time of $t_w$, a postsurgery length of stay of $t$, and a discharge to destination $r$, conditional upon both measured and unmeasured characteristics. Estimation of the model requires us to integrate out the unmeasured characteristics $\upsilon$. Denoting individuals by $i$ and hospitals by $h$, let the indicator variable $\delta_{ihs}$ ($\delta_{ihn}$) equal 1 if individual $i$ in hospital $h$ undergoes (does not undergo) surgery, and 0 otherwise, and $\delta_{ihr}$ is defined similarly if individual $i$ in hospital $h$ is discharged to destination $r$ after surgery. If the length-of-stay transition intensities are allowed to depend upon $\upsilon$, the likelihood function for the model is then given by

$$L = \prod_i \int f_n(t_{ihw}|X_{ihw}, \upsilon)^{\delta_{ihn}} \times \left\{ f_s(t_{ihw}|X_{ihw}, \upsilon)\prod_{r\in D} f_r(t_{ih}|X_{ih}, \upsilon)^{\delta_{ihr}} \right\}^{\delta_{ihs}} dG(\upsilon) \tag{4}$$

The first term of equation (4) is the probability of observing a stay of $t_{ihw}$ days that does not result in surgery for individual $i$, while the term in braces is the joint probability of observing a wait of $t_{ihw}$ days prior to receiving surgery and a postsurgical length of stay of $t_{ih}$ days with release to destination $r$.

### B. *Specification of Transition Intensities*

The next step in the construction of the empirical model involves the specification of the functional form of the transition intensities in equation (4). We follow the common approach in the literature (Kiefer (1988), Lancaster (1990)) and use a proportional-hazards specification. In addition, we allow the unmeasured correlated component $\upsilon$ to have different factor loadings $\theta_r$[10] in each transition-intensity function, so that

$$\lambda_r(t_{ih}|X_{ih}, \upsilon) = \exp(X_{ih}\beta_r + \gamma_{hr} + \theta_r\upsilon)\lambda_{0r}(t_{ih}) \qquad r = s, n, a, d, \tag{5}$$

where $\lambda_{0r}(t)$ represents the baseline transition-intensity function. The indicator variable $\gamma_{hr}$ equals 1 if the patient was admitted to hospital $h$, and 0 otherwise. Inclusion of the full set of hospital dummies in each transition intensity allows us to distinguish between hospital and patient factors in explaining wait times and outcomes. For example, patients admitted to inefficient hospitals may be delayed longer and have poorer outcomes. Consequently, using the hospital fixed-effects specification, the impact of patient-level characteristics $X_{ih}$ is identified by within-hospital variation.

A variety of parametric and nonparametric methods are available to estimate the baseline transition intensity (Lancaster (1990)). Some guidance as to the appropriate specification may be gained by examining the empirical transition intensities for live discharge and inhospital mortality, as well as the conditional probabilities of surgery and no surgery. Figures 1(a) through (d) show that the live discharge and surgery transition intensities have an inverted U shape in both countries, while the other transition intensities are fairly constant. After estimating a variety of specifications, we found that a parsimonious specification of the baseline transition intensity that allows for nonmonotonic behavior of the type shown in the figures and that yields a reasonable fit of the data is the log-logistic distribution:

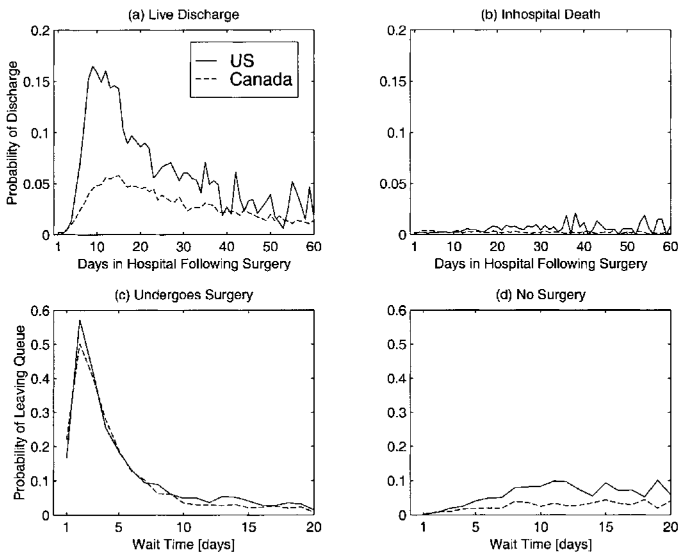
$$\lambda_{0r}(t) = \frac{\rho_r\alpha_r t^{\alpha_r - 1}}{1 + \rho_r t^{\alpha_r}} \qquad \alpha_r > 0, \rho_r > 0 \qquad r = s, n, a, d. \tag{6}$$

Note that specification (6) allows the parameters of the baseline transition intensities to differ for each destination.

[9] It may be argued that not all admitted hip-fracture patients are at risk of undergoing surgery. However, the presence of a not-at-risk subpopulation may be captured by the unobserved heterogeneity distribution (Pickles and Crouchley (1995)).

[10] Not all of the factor loadings are identified. We normalize $\theta_s = 1$.

FIGURE 1.—EMPIRICAL TRANSITION INTENSITIES




In order to integrate out the unmeasured scalar heterogeneity term in the likelihood function, we must choose a functional form for the distribution of $\upsilon$, $G(\upsilon)$. We adopt the semiparametric approach suggested by Heckman and Singer (1984) and assume that $G(\upsilon)$ may be approximated by a discrete distribution with a finite number of values, which are the points of support for the distribution of $\upsilon$. This approach has been applied in other health care contexts (Cutler (1995); Goldman (1995); Hamilton, Hamilton, and Mayo (1996)), and is robust to potential misspecification of the functional form of the heterogeneity distribution (such as gamma distribution) which may lead to biased estimates of the parameters (Heckman and Singer (1984), Lancaster (1990)). The location of the points of support and their associated probabilities are estimated jointly with the other parameters of the model. With this specification of $G(\upsilon)$, the likelihood function may be written as

$$L(\beta, \gamma, \alpha, \rho, \omega, \theta, \upsilon)$$
$$= \prod_i \sum_{k=1}^{K} \omega_k f_n(t_{ihw}|X_{ihw}, \upsilon_k)^{\delta_{ihn}} \times \left\{ f_s(t_{ihw}|X_{ihw}, \upsilon_k) \prod_{r \in D} f_r(t_{ih}|X_{ih}, \upsilon_k)^{\delta_{ihr}} \right\}^{\delta_{ihs}}. \tag{7}$$

where $\upsilon_k$, $k = 1, \ldots, K$ are the points of support with associated probabilities $\omega_k$ which sum to 1. Empirical studies in the literature suggest that the value of $K$ required for an adequate nonparametric representation of $G(\omega)$ is small in most cases, usually on the order of $K = 2$ or 3.[11]

[11] Lindsay (1983) presents the conditions under which more points of support are necessary.

## C. Explanatory Variables

We estimate equation (7) separately for the Quebec and Massachusetts data sets using maximum likelihood. As noted above, between-hospital differences in wait times and outcomes (within each country) are captured by the hospital fixed effects.[12] We are particularly interested in the impact of wait time for surgery, $t_w$, on postsurgery outcomes, because this tells us the impact of queuing on outcomes. We are also interested in whether lower income and race (in the United States) correlates with longer delays and poorer outcomes. For Quebec patients, income is measured by median male income in 1988 in the postal code of residence. For Massachusetts, the NIS data set may be used to categorize patients into rough income quartiles based upon average household income in their postal code of residence. We then use the Quebec data to create income quartiles analogous to those for Massachusetts.[13] Thus, dummy variables for income quartiles are included in both regressions, with the highest-income quartile as the excluded category. A dummy variable for white versus nonwhite patients is also included in the Massachusetts analyses. Another important set of covariates are dummy indicators for the comorbidities that comprise the Charlson comorbidity index. Indicator variables for the ten comorbidities that comprise the index are included in the regressions.[14] Other patient-specific covariates include age, gender, and dummy variables for pertrochanteric fractures and other/unspecified fractures relative to the excluded category of transcervical fractures, as well as indicators for year of admission.[15]

Delay until surgery is an endogenous variable in our model that depends on covariates $X_w$, but it is also the key covariate included in $X$ to explain surgical outcomes (mortality and length of stay). To help identify the model, the vector $X_w$ should contain at least one predictor not included in $X$. This is a stringent requirement, because a variable is needed that predicts delays but not health status (and therefore postsurgical outcomes). A natural candidate is the day of the week of admission to the hospital. Surgical staff prefer to operate on weekdays rather than on weekends, and it appears that this preference systematically affects the distribution of wait times across days of the week. Column 1 of table 2 shows that admissions are roughly equally distributed across

[12] Descriptive statistics on these variables for both countries are included in table A1.

[13] Postal code income is reported in categories in the United States. 27% of Massachusetts come from postal-code areas with average family incomes below \$30,000; 20% between \$30,000 and \$35,000; 33% between \$35,000 and \$45,000; and 20% above \$45,000. We construct dummy variables for these categories. For Quebec, we construct analogous income categories: 27% of patients come from postal codes with median male incomes below \$18,500; 20% between \$18,500 and \$20,400; 33% between \$20,400 and \$25,300; and 20% above \$25,300.

[14] These comorbidities are myocardial infarction, peripheral vascular disease, dementia, chronic pulmonary disease, rheumatologic disease, liver disease, diabetes, renal disease, malignancy, and metastatic solid tumor.

[15] In the subsequent tables, the coefficient estimates of these variables are not surprising and hence not reported in the interest of conserving space.

TABLE 2.—RELATIONSHIP BETWEEN ADMITTING DAY, WAIT TIME, CASE MIX, AND OUTCOMES FOR PATIENTS UNDERGOING SURGERY, BY COUNTRY

| Admitting Day | Frac. | Fraction with Delay < 3 | Charlson Index | Fraction Died in Hospital | Postsurgery Length of Stay (Mean) |
|---|---|---|---|---|---|
| Panel A: Canada (Quebec) | | | | | |
| Monday | 0.14 | 0.62 | 0.54 | 0.067 | 28.0 |
| Tuesday | 0.14 | 0.64 | 0.57 | 0.084 | 27.0 |
| Wednesday | 0.15 | 0.64 | 0.59 | 0.087 | 27.7 |
| Thursday | 0.15 | 0.73 | 0.51 | 0.084 | 27.4 |
| Friday | 0.15 | 0.69 | 0.49 | 0.071 | 27.9 |
| Saturday | 0.14 | 0.66 | 0.50 | 0.078 | 29.7 |
| Sunday | 0.13 | 0.66 | 0.55 | 0.077 | 28.7 |
| $p$-value[1] | — | 0.000 | 0.108 | 0.340 | 0.651 |
| Panel B: United States (Massachusetts) | | | | | |
| Monday | 0.15 | 0.69 | 0.64 | 0.029 | 15.5 |
| Tuesday | 0.15 | 0.70 | 0.66 | 0.049 | 13.7 |
| Wednesday | 0.14 | 0.72 | 0.63 | 0.043 | 14.0 |
| Thursday | 0.14 | 0.76 | 0.59 | 0.047 | 13.4 |
| Friday | 0.15 | 0.67 | 0.58 | 0.043 | 13.9 |
| Saturday | 0.14 | 0.66 | 0.58 | 0.046 | 13.7 |
| Sunday | 0.13 | 0.70 | 0.61 | 0.045 | 13.1 |
| $p$-value[1] | | 0.000 | 0.372 | 0.269 | 0.126 |

[1] $p$-value from test of the null hypothesis of equality of means across day of week of admittance.

days of the week in both Canada and the United States. However, the bottom row of column 2 indicates that the proportion of patients undergoing surgery within two days of admittance differs significantly across days of the week in both countries. No significant differences were found in the Charlson comorbidity index, age, income, or other characteristics across patients admitted on different days. Therefore, it is unlikely that the day of the week of admission is correlated with health status. Consequently, we include dummy indicators for each day as explanators in the wait-time transition intensities, but exclude them from the length of stay equations. Moreover, columns 4 and 5 of table 2 show that, despite differences in surgery wait times across days of the week, little difference is observed in postsurgical outcomes by day of the week of admission. This table provides prima facie evidence that wait time is likely to have only a small effect, if any, on outcomes.

## IV. Results

Tables 3A and 3B present the parameter estimates for the U.S. and Canadian models, respectively. Positive coefficients indicate that an increase in the variable is associated with an increase in the associated transition intensity. We found that the heterogeneity distribution $G(\upsilon)$ can be approximated by a finite distribution with two points of support ($K = 2$) in both Canada and the United States.

TABLE 3A.—U.S. PROPORTIONAL HAZARD ESTIMATES BASELINE HAZARD SPECIFICATION: LOG-LOGISTIC

| | Surgery Decision | | Postsurgery LOS | |
|---|---|---|---|---|
| | Surgery (s) | No Surgery (n) | Discharged Alive | Discharged Dead |
| Dementia | −0.045 | −0.175 | −0.174 | −0.568 |
| | (−0.910) | (−1.028) | (−4.920) | (−2.485) |
| Pulmonary disease | −0.089 | 0.053 | 0.029 | 0.511 |
| | (−2.132) | (0.419) | (0.831) | (3.452) |
| Diabetes | −0.088 | 0.051 | 0.028 | 0.041 |
| | (−1.836) | (0.353) | (0.745) | (0.204) |
| Liver disease | −0.608 | −0.161 | −0.646 | 1.100 |
| | (−2.725) | (−0.354) | (−2.623) | (2.112) |
| Monday | −0.251 | 0.219 | | |
| | (−4.703) | (1.445) | | |
| Tuesday | −0.209 | 0.046 | | |
| | (−3.899) | (0.306) | | |
| Wednesday | −0.147 | 0.116 | | |
| | (−2.685) | (0.755) | | |
| Friday | −0.307 | −0.043 | | |
| | (−5.856) | (−0.270) | | |
| Saturday | −0.296 | 0.064 | | |
| | (−5.384) | (0.424) | | |
| Sunday | −0.274 | 0.094 | | |
| | (−5.023) | (0.601) | | |
| Income Q1 | −0.127 | −0.027 | −0.117 | 0.102 |
| | (−2.237) | (−0.168) | (−2.622) | (0.371) |
| Income Q2 | −0.135 | −0.118 | −0.067 | 0.231 |
| | (−2.288) | (−0.734) | (−1.494) | (0.840) |
| Income Q3 | −0.097 | −0.054 | −0.034 | 0.374 |
| | (−1.930) | (−0.377) | (−0.906) | (1.564) |
| White | 0.115 | 0.088 | 0.080 | 0.332 |
| | (1.822) | (0.495) | (1.482) | (1.048) |
| $t_w$ | | | −0.008 | −0.025 |
| | | | (−1.206) | (−0.677) |
| $\rho$ | 0.027 | 0.0002 | 0.00002 | 0.0002 |
| | (4.717) | (2.082) | (6.803) | (3.215) |
| $\alpha$ | 3.983 | 2.555 | 5.212 | 1.576 |
| | (45.457) | (59.200) | (58.824) | (18.182) |
| $\theta$ | 1 | −0.700 | 0.151 | −0.245 |
| | | (−2.662) | (5.568) | (−1.971) |

NOTE: $t$-statistics in parentheses. Models also contain constants, hospital indicators, and controls for age, gender, year dummies, indicators for pertrochanteric and other fractures, and comorbidity dummies for myocardial infarction, peripheral vascular disease, rheumatologic disease, liver disease, renal disease, malignancy, and metastatic solid tumor.

### A. *Wait Times*

The first two columns in tables 3A and 3B present the wait-time results for surgery (column 1) and no surgery (column 2). The coefficients on the income dummies in the United States imply that patients living in higher-income areas tend to leave the queue for hip-fracture surgery more quickly than lower-income persons in a given hospital. In contrast, income does not explain the variation in wait times in Canada. Thus, it would appear that higher-income persons receive preferential treatment in the United States, but not in Canada, which has universal first-dollar health insurance coverage. Of course, it may still be the case that poorer individuals tend to be admitted to "slower" hospitals in Canada. We investigate these hospital-level effects below.[16]

The presence of comorbidities lead to lower probabilities of both leaving the queue for surgery and for leaving the hospital without surgery. The day-of-the-week variables are

[16] Because we proxy for household income using postal-code income, our estimates will be biased toward zero.

TABLE 3B.—CANADA PROPORTIONAL HAZARD ESTIMATES
BASELINE HAZARD SPECIFICATION: LOG-LOGISTIC

| | Surgery Decision | | Postsurgery LOS | |
|---|---|---|---|---|
| | Surgery (s) | No Surgery (n) | Discharged Alive | Discharged Dead |
| Dementia | −0.107 | −0.282 | −0.361 | 0.254 |
| | (−2.187) | (−1.866) | (−9.278) | (2.392) |
| Pulmonary disease | −0.103 | −0.272 | −0.194 | 0.450 |
| | (−2.584) | (−2.457) | (−5.073) | (4.839) |
| Diabetes | −0.035 | −0.276 | −0.246 | 0.223 |
| | (−0.845) | (−1.900) | (−6.349) | (2.067) |
| Liver disease | −0.245 | −0.356 | −0.471 | 1.048 |
| | (−1.703) | (−1.000) | (−2.690) | (4.378) |
| Monday | −0.193 | 0.198 | | |
| | (−4.527) | (1.487) | | |
| Tuesday | −0.150 | −0.0008 | | |
| | (−3.537) | (−0.007) | | |
| Wednesday | −0.054 | −0.014 | | |
| | (−1.266) | (−0.099) | | |
| Friday | −0.126 | 0.130 | | |
| | (−3.019) | (1.000) | | |
| Saturday | −0.070 | 0.209 | | |
| | (−1.584) | (1.360) | | |
| Sunday | −0.058 | 0.287 | | |
| | (−1.354) | (2.013) | | |
| Income Q1 | −0.026 | 0.145 | 0.121 | −0.119 |
| | (−0.662) | (1.074) | (3.499) | (−0.928) |
| Income Q2 | −0.030 | 0.078 | 0.071 | −0.079 |
| | (−0.714) | (0.550) | (1.857) | (−0.591) |
| Income Q3 | −0.029 | 0.039 | 0.022 | 0.111 |
| | (−0.791) | (0.329) | (0.674) | (0.985) |
| $t_w$ | | | −0.002 | −0.002 |
| | | | (−0.632) | (−0.568) |
| $\rho$ | 0.030 | 0.00001 | 0.0005 | 0.012 |
| | (5.236) | (1.344) | (10.638) | (6.936) |
| $\alpha$ | 3.290 | 4.121 | 2.962 | 1.290 |
| | (35.714) | (15.150) | (58.831) | (15.873) |
| $\theta$ | 1 | −0.170 | 0.225 | −0.136 |
| | | (−1.487) | (9.228) | (−2.130) |

Note: $t$-statistics in parentheses. Models also contain constants, hospital indicators, and controls for age, gender, year dummies, indicators for pertrochanteric and other fractures, and comorbidity dummies for myocardial infarction, peripheral vascular disease, rheumatologic disease, liver disease, renal disease, malignancy, and metastatic solid tumor.

jointly significant,[17] with patients admitted on Thursdays receiving surgery faster than patients admitted on other days of the week in both Quebec and Massachusetts. This pattern suggests a strong effort to complete surgery before the weekend.

### B. *Length of Stay*

The third column of tables 3A and 3B relates to postsurgery length of stay inhospital (conditional upon being discharged alive). Our primary interest is the effect of wait time. Wait times do not affect the hazard of a live discharge in either country, suggesting that the relationship between delay and postsurgical outcomes observed in table 1 reflects unmeasured patient- or hospital-level factors. It is interesting to note that, if the same specification is reestimated without unobserved heterogeneity, the point estimate of $t_w$ increases in magnitude to $-0.039$ ($t = -6.674$) in Massachusetts and to $-0.016$ ($t = -6.959$) in Quebec.[18] Thus, the strong statistical relationship between delay and outcomes observed in table 1 (as well as in many previous studies) appears to reflect unmeasured patient frailty at the time of admission.

The factor loading, $\theta_a$, is positive and precisely estimated in both countries, so unobserved factors leading to shorter wait times for surgery also increase the conditional probability of being discharged alive. These findings are consistent with the hypothesis that the heterogeneity factor $\upsilon$ proxies for patient frailty at the time of admission to hospital. Lower-income persons in Massachusetts have a lower probability of being discharged alive at each point in time relative to patients in the highest-income quartile. In contrast, the results in column 3 of table 3B imply that patients in the lowest-income quartile have a higher probability of being discharged alive at any time $t$ in Quebec than do individuals in the highest-income quartile. We will assess the magnitude of this effect with simulations conducted below.

### C. *Mortality*

The final column in tables 3A and 3B displays the parameter estimates for the hazard of being discharged dead. In both countries, the coefficient estimate on $t_w$ is not precisely estimated, indicating that delay does not influence the hazard for inpatient death. Income also does not appear to affect the probability of death in hospital in either country. Not surprisingly, comorbidities tend to increase the hazard of being discharged dead.

### D. *Heterogeneity*

Table 3C presents the estimates of the unobserved heterogeneity distributions. The data indicate clustering around two points, $\upsilon_1$ and $\upsilon_2$, in both Canada and the United States, with $\upsilon_1 < \upsilon_2$. For convenience, denote a realization of $\upsilon_k$ as a "type $k$" patient. One interpretation of the $K = 2$ support points is that there are two types of patients in Canada and the United States. Approximately 13% of the Canadian

[17] A likelihood-ratio test of the null hypothesis that the coefficients on the day-of-the-week variables are jointly zero in the surgery and no-surgery transition intensities has a $p$-value of 0.005 in each country. A likelihood-ratio test of the null hypothesis that the day-of-week variables do not enter the live and dead discharge transition intensities cannot be rejected ($p$-value = 0.107 in the United States and $p$-value = 0.533 in Canada). Moreover, the coefficient estimates on the other independent variables in the live and dead discharge transition intensities remain unchanged when day of week is included. These tests support the exclusion restriction on the day-of-the-week variables.

[18] The estimated coefficients on the other variables in the model without unobserved heterogeneity are nearly identical to those reported in tables 3A and 3B in almost all cases. In addition, there may be concern that misspecification of the baseline hazard as log-logistic biases the parameter estimates. However, we reestimated each of the transition intensities in each country (in the no-unobserved-heterogeneity case) using a partial-likelihood approach that allows for an arbitrary form of the baseline hazard (Kalbfleisch and Prentice (1980)). Again, we found no significant difference in the estimated coefficients using the two approaches. These results are available from the authors.

TABLE 3C.—PROPORTIONAL HAZARD MODEL
HETEROGENEITY PARAMETER ESTIMATES
BASELINE HAZARD SPECIFICATION: LOG-LOGISTIC

| | United States | Canada |
|---|---|---|
| $\upsilon_1$ | −1.696 | −1.755 |
| $\upsilon_2$ | 0.956 | 1.153 |
| | (8.365) | (13.888) |
| $\omega_1$ | 0.179 | 0.127 |
| | (9.089) | (4.513) |
| Log-likelihood | −43942.0 | −71995.8 |
| $N$ | 8979 | 12016 |

Note: $t$-statistics in parentheses. $t$-statistic on $\upsilon_2$ is from a test of the null hypothesis that $\upsilon_1 = \upsilon_2$.

sample are type 1 individuals who experience significantly longer delays for surgery, are less likely to undergo surgery, and have longer postsurgery lengths of stay and higher probabilities of inhospital mortality than the other patient types. Similar statements can be made about the United States, where the data indicates that there are two types of patients, with type 1 patients (18% of the sample) experiencing longer delays and prolonged hospital stays when compared to type 2 individuals.

### E. *Patient versus Hospital Effects*

The coefficient estimates presented in table 3 incorporate hospital fixed effects, and thus are identified by within-hospital variation. However, it still may be the case that many of the factors of interest (such as wait time and income) operate at the hospital rather than patient level. For example, hospitals that tend to have longer delays may have poorer postsurgical outcomes, and low-income patients may tend to be admitted to inefficient hospitals. To assess the contribution of hospital-specific effects, we reestimated the models in table 3 for each country excluding the hospital dummies, so that the coefficients on the patient characteristics reflect both within- and between-hospital differences.

Table 4 presents the coefficient estimates for income, race, and wait time when the hospital indicators are excluded from the model. In the case of the United States, comparison of table 3A and panel A of table 4 shows two notable differences. First, the coefficients on the income dummies in the live discharge transition intensity almost triple in magnitude, implying that low-income individuals tend to be admitted to hospitals that have lower conditional probabilities of live discharge. Second, wait time has a negative and significant impact on live discharge when hospital dummies are excluded, suggesting that hospitals that tend to have longer delays for all their patients also have poorer outcomes.

The Canadian results in panel B of table 4 also imply that low-income patients tend to be admitted to hospitals with relatively long wait times, although table 3B showed that, conditional upon hospital, there was no income effect. In addition, as in the United States, the coefficient on wait time in the live discharge equation triples when hospital indicators are excluded, implying that Canadian hospitals with longer wait times also tend to have longer lengths of stay.

TABLE 4.—PROPORTIONAL HAZARD ESTIMATES EXCLUDING HOSPITAL INDICATORS
BASELINE HAZARD SPECIFICATION: LOG-LOGISTIC

| | Surgery Decision | | Postsurgery LOS | |
|---|---|---|---|---|
| | Surgery (s) | No Surgery (n) | Discharged Alive | Discharged Dead |
| **Panel A: United States** | | | | |
| Income Q1 | −0.168 | −0.203 | −0.303 | −0.197 |
| | (−3.511) | (−1.748) | (−8.440) | (−0.974) |
| Income Q2 | −0.017 | −0.184 | −0.189 | −0.061 |
| | (−0.342) | (−1.508) | (−5.272) | (−0.300) |
| Income Q3 | −0.075 | −0.166 | −0.130 | 0.115 |
| | (−1.704) | (−1.475) | (−4.109) | (0.640) |
| White | 0.038 | 0.013 | 0.125 | 0.338 |
| | (0.601) | (0.084) | (2.505) | (1.255) |
| $t_w$ | | | −0.018 | −0.048 |
| | | | (−2.589) | (−1.458) |
| **Panel B: Canada** | | | | |
| Income Q1 | −0.132 | −0.180 | 0.103 | −0.180 |
| | (−3.226) | (−1.467) | (3.525) | (−1.707) |
| Income Q2 | −0.024 | 0.099 | 0.051 | −0.042 |
| | (−0.534) | (0.746) | (1.564) | (−0.392) |
| Income Q3 | −0.029 | 0.015 | 0.047 | 0.061 |
| | (−0.723) | (0.134) | (1.673) | (0.626) |
| $t_w$ | | | −0.006 | 0.0005 |
| | | | (−1.652) | (0.195) |

Note: $t$-statistics in parentheses. Model specification the same as in table 3, except hospital indicators are excluded.

## V. Simulations

The parameter estimates show the direction and significance of the variables in the model, but one cannot easily interpret the magnitude of their impact. Thus, we construct the unconditional (on time) probability of leaving the hospital for destination $r$ (denoted $P_r$) for a variety of patient types,[19] as well as the unconditional (on destination) length of hospital stay after surgery, $E_t$.[20] In addition, we calculate the unconditional probability of receiving surgery, $P_s$, as well as the expected delay for individuals receiving surgery, $E_s$. Two sets of simulations are conducted. The first set examines the impact of patient characteristics on wait times and outcomes assuming the patient is admitted to the average hospital in each country. Rather than pick one hospital in each country to make comparisons, the simulations are conducted using estimates from a model that excludes the hospital dummies. However, to account for

[19] The unconditional probability of leaving the hospital for reason $r$ is given by

$$P_r = \int_0^\infty \lambda_r(t) \prod_{j \in D} \exp\left(-\int_0^t \lambda_j(u)\, du\right) dt.$$

The exponential terms on the right side of the equation represent the probability that the individual does not leave the hospital before period $t$ for any destination.

[20] The expected postsurgery length of stay conditional upon exit destination $r$ is given by

$$E_r = (P_r)^{-1} \int_0^\infty t\lambda_r(t) \prod_{j \in D} \exp\left(-\int_0^t \lambda_j(u)\, du\right) dt.$$

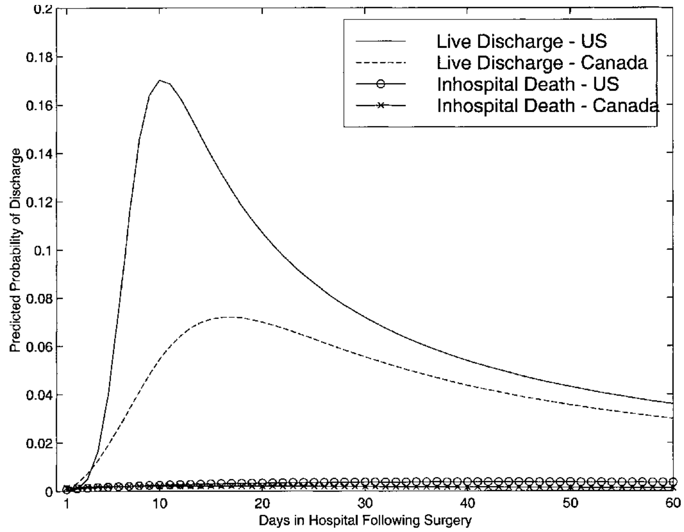
The unconditional (on destination) expected postsurgery length of stay is then $E_t = \Sigma_r P_r E_r$.

TABLE 5.—PREDICTED DESTINATION PROBABILITIES AND EXPECTED WAIT TIMES AND LENGTH OF STAY, BY PATIENT TYPE AND COUNTRY

| Patient Type | Prob. Of Surgery $P_s$ | Mean Wait Time $E_s$ | Prob. Of Live Discharge $P_a$ | Prob. Of In-hospital Death $P_d$ | Mean Length of Stay $E_t$ |
|---|---|---|---|---|---|
| Panel A: Canada | | | | | |
| Base[1] | 0.944 | 3.14 | 0.955 | 0.045 | 24.4 |
| Severe Comorbidities[2] | 0.940 | 4.77 | 0.727 | 0.273 | 48.7 |
| Income Q1 | 0.944 | 3.54 | 0.962 | 0.038 | 23.4 |
| Income Q4 | 0.948 | 3.19 | 0.951 | 0.049 | 25.6 |
| $t_w$ = 1 | — | — | 0.956 | 0.044 | 24.1 |
| $t_w$ = 7 | — | — | 0.954 | 0.046 | 24.9 |
| Panel B: United States | | | | | |
| Base[1] | 0.949 | 2.44 | 0.971 | 0.029 | 13.3 |
| Severe Comorbidities[2] | 0.894 | 3.80 | 0.873 | 0.127 | 29.0 |
| Income Q1 | 0.943 | 2.59 | 0.971 | 0.029 | 14.7 |
| Income Q4 | 0.945 | 2.38 | 0.976 | 0.024 | 11.5 |
| Nonwhite | 0.948 | 2.48 | 0.975 | 0.025 | 14.9 |
| $t_w$ = 1 | — | — | 0.970 | 0.030 | 13.0 |
| $t_w$ = 7 | — | — | 0.975 | 0.025 | 14.3 |

[1] Base patient is 78 years old, female, has income in the second quartile, has 0 comorbidities, a pertrochanteric fracture, admitted to a teaching hospital performing 54 surgeries in 1991 on a Thursday.
[2] Severe comorbidities include dementia and liver disease.

FIGURE 2.—PREDICTED TRANSITION INTENSITIES, BASE PATIENT




differences in hospital volumes and teaching status across countries, these variables are included as controls in the model. $P_s$, $E_s$, $P_r$, and $E_t$, are calculated for a base patient type with a specified set of characteristics. We then simulate the changes in these variables associated with changes in certain characteristics of the base patient. The second set of simulations investigates the magnitude of hospital differences in each country by constructing hospital-specific predicted outcomes for the base patient in each country.

The first set of simulations appear in table 5. The base case is a 78-year-old female patient who has no comorbidities, a pertrochanteric fracture, and who is admitted on a Thursday in 1991 to a teaching hospital performing 54 surgeries that year. The income of this baseline patient lies in the second quartile in each country. How would this patient fare in each country? In both locations, she has about a 95% chance of probability of receiving surgery. However, she will have to wait 29% longer in Canada (3.14 days versus 2.44 days). She is also less likely to die in the United States, and will be discharged much more quickly if she lives.

We also run simulations changing income quartiles. Tables 3 and 4 indicated that higher-income persons exit the surgery queue faster than lower-income patients. However, there is little difference across income quartiles in the probability of surgery and/or wait times in either country. Thus, although the income effect is precisely estimated in our model, its quantitative impact on surgery queues is small. In contrast, comorbidities appear to have a much larger impact on surgery queues. In both countries, if our base patient were admitted with ''Severe Comorbidities'' (dementia and liver disease), then she would wait 50% longer for surgery.

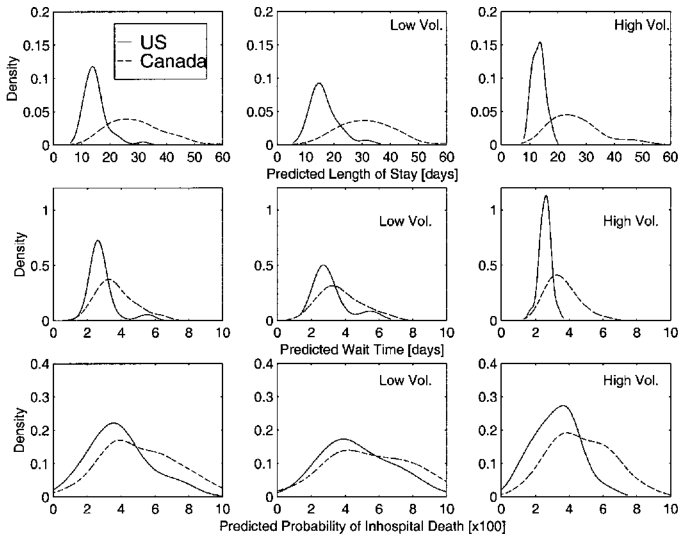
Finally, if we alter our base patient's wait time for surgery to be either one day or seven days, the differential postsurgery effects are surprisingly modest. The only substantial difference associated with an extended delay is that our patient has an expected length of stay (postsurgery) that is one day longer in both countries. Given that we have chosen a relatively wide difference in delay times to conduct these simulations, these results imply that the effect of queuing for surgery on postsurgery outcomes is relatively inconsequential.

One strong pattern emerges from table 5. Canada has much higher expected lengths of stay and probabilities of inpatient mortality for every simulation. Yet this difference is not explained by surgical delays, according to our simulations. A further examination of the descriptive statistics in table A1 indicates that differences in other patient or hospital characteristics do not appear to explain these results either. However, as mentioned previously, longer postsurgery lengths of stay in Canada may contribute to the higher inpatient mortality rates. To test this hypothesis, we used the estimates derived from our model to predict the conditional transition intensities for live and dead discharge for the base patient. These predictions are presented in figure 2. The predicted live discharge transition intensity for the United States lies substantially above that for Canada, while the predicted hazard for inpatient mortality is slightly higher in the United States. Thus, the higher Canadian inpatient mortality rates observed in the raw data in table 1 and in the simulations in table 4 are due to the longer lengths of stay in the Canadian patient population. With respect to hip-fracture patients, the Canadian health care system exhibits longer postsurgery lengths of stay, but treatment in this country does not lead to increased mortality relative to the United States.

### A. *Assessing the Importance of Hospital Effects*

In the second set of simulations, predicted wait times and outcomes for the base patient were generated for each of the

FIGURE 3.—DISTRIBUTIONS OF HOSPITAL-SPECIFIC OUTCOMES: OVERALL, AND BY VOLUME.




68 Quebec and 45 Massachusetts hospitals. The distributions of these predicted quantities are presented in figure 3. The most notable feature of the figure is that the distribution of hospital average postsurgery length of stay in the United States is much more concentrated than in Canada. The top-left graph indicates that hospital averages range from roughly 8 to 36 days in the United States, compared with 8 to 60 days in Canada. The graphs in the second and third columns indicate that this does not reflect volume differences across countries, since even high-volume hospitals (hospitals performing more than 54 surgeries per year) in Canada exhibit substantial variation. Note, however, that the modes of the distributions indicate that low-volume hospitals in Canada tend to have mean lengths of stay that are approximately one week longer than those at high-volume Canadian hospitals. In contrast, the modes of the U.S. distributions are roughly the same for low- and high-volume hospitals, suggesting that volume effects are more important in Canada. In the case of wait times, the graphs in the second row of the figure indicate greater hospital variation in Canada than the United States, although the country differences are less substantial than those observed for length of stay. Finally, similar international differences are found for predicted mean probabilities of inhospital mortality among high-volume hospitals. Consequently, not only does the base patient experience a substantially shorter delay and postsurgery length of stay in the United States, but hospital-specific variation in outcomes is substantially higher in Canada.

## VI. Conclusions

Critics of the Canadian health care system have argued that universal health insurance with constrained resource allocation leads to queuing for health care services, which can have detrimental health effects. However, relatively little research has been conducted to assess the determinants of queuing, and to assess how these queues affect outcomes across countries. We consider this argument by looking at queuing for hip-fracture surgery, a common injury. We find that longer delays lead to significantly longer stays in the hospital after surgery, but they do not affect mortality, although the former result reflects hospital-level differences. Further, our simulations suggest the magnitude of any effects are quite modest. Thus, although surgery delays are slightly longer in Canada, these delays do not have a substantive impact on postsurgery outcomes in either country.

The databases we used did not follow patients after discharge, so that we were unable to compare postdischarge morbidity or mortality in Canada and the United States. However, our initial goal was to test for a relationship between surgery delay and patient outcomes, and in-hospital data is the best data to examine this issue. In fact, past studies caution against the use of longer-term data to assess quality or the effectiveness of surgical treatment. Longer-term postdischarge data commonly corroborates patient-outcomes data collected within the first thirty days of admission. And, in cases where it does not, long-term outcomes potentially are a function of postdischarge outpatient care (Garnick, DeLong, and Luft (1995)). In addition, longer-term data for hip-fracture patients is likely to reflect the underlying frailty of these individuals prior to admission, rather than the quality of hospital care (Seagroatt and Goldacre (1994)).

We originally hypothesized that outcomes would also be more equitable in Canada, with its system of universal health insurance. We find that higher-income persons in both Quebec and Massachusetts exit the surgery queue more quickly than lower-income patients. In Quebec, this differential reflects differential access to hospitals across income classes, while, in Massachusetts, wait times differ by income within a given hospital. Still, our simulations indicate that differences across income classes in postsurgery length of stay are wider in the United States. Higher-income patients also have a higher probability of being discharged alive on any day postsurgery in the United States. One might attribute the slower discharge of low-income patients to inadequate postsurgery care; however, it may be that there are fewer outpatient therapeutic options (such as home health care and nursing homes) available to the poor. Thus, physicians keep them in the hospital longer. Further investigation seems warranted.

Nevertheless, the different outcomes by income class are overshadowed by the dramatic international differences in length of stay. Controlling for patient and hospital characteristics, postsurgery length of stay in Quebec is almost double that observed in Massachusetts. Differences in observable characteristics between the two countries do not explain the differential. Longer inpatient hospital stays in Canada versus the United States have been noted by previous researchers, and are likely due to different reimbursement mechanisms in the two countries. U.S. hospitals are reimbursed a fixed price for each admission based on the DRG system, which encourages prompt discharge. In fact, past research has noted a 42% decline in length of stay for patients treated in a large U.S. community hospital after the introduction of DRG reimbursement in the mid-1980s (Fitzgerald, Moore, and Dittus (1988)). Even though patients were being discharged much earlier, one-year mortality did not increase; although the proportion of patients discharged to nursing homes rose.

In contrast, Canadian hospitals receive a global budget each year, which is not directly related to patient length of stay. Thus, the incentive to promptly transfer patients from acute-care settings following surgery is not as strong in Canada. Our model also shows more variability in delays and length of stay across hospitals in Canada. Perhaps DRG reimbursement may have forced all U.S. hospitals to become efficient in terms of discharging patients, while such incentives do not exist for Canadian facilities. Reductions in lengths of stay in Canada among inefficient hospitals would free up beds for those patients still waiting for elective inpatient surgery, or permit cost savings through the closure of unneeded beds. Even if reduced lengths of stay lead to greater demand for nursing homes and rehabilitation facilities, these institutions can provide appropriate care for recuperating patients at a lower cost than acute-care hospitals.

The above discussion indicates that further examination of Canadian hospital-specific variations has important implications for patient welfare and health care costs. For example, the poorer outcomes associated with low-volume hospitals suggests that hospital consolidation should be considered in Canada. The population in Quebec is more widely dispersed than in Massachusetts, which may require a greater number of relatively small hospitals to serve more rural locations. However, the wider population dispersion cannot entirely justify the lower surgical volume observed in the data. For instance, the island of Montreal, an area of only 176 square kilometers, has 39 operating hospitals. In contrast, the city of Boston covers 125 square kilometers and has 25 hospitals.

Finally, we find, as others have, that Canada has a higher inpatient mortality rate than the United States. By estimating our model using a competing-risks approach, we demonstrate that this phenomenon can be attributed to the longer lengths of stay in Quebec, rather than to any systemic difference in outcomes on any postsurgery day. This finding should resolve some of the confusion over the differential mortality rates. All these results, taken together, suggest that intense focus on queuing for health care services may be misdirected. Wider differences in hospital-specific treatment and outcomes in Canada versus the United States may have more-important implications for patients' health. Detailed comparisons of health care treatment and outcomes in other illnesses is required to confirm these findings.

## REFERENCES


Baker, S. P., B. O'Neill, and R. S. Karpf, *The Injury Fact Book* (Lexington, MA: D. C. Heath and Company, 1984).

Bredahl, C., B. Nyholm, K. B., Hindsholm, J. S. Mortensen, and A. S. Oleson, ''Mortality after Hip Fracture: Results of Operation Within 12 h of Admission,'' *Injury* 23 (1992), 83–86.

Charlson, M. E., P. Pompei, K. Ales, and C. R. MacKenzie, ''A New Method of Classifying Prognostic Comorbidity in Longitudinal Studies: Development and Validation,'' *Journal of Chronic Diseases* 40 (1987), 373–383.

Cutler, D., ''The Incidence of Adverse Medical Outcomes Under Prospective Payment,'' *Econometrica* 63 (1995), 29–50.

Davis, F. M., D. F. Woolner, C. Frampton, A. Wilkinson, A. Grant, R. T. Harrison, M. T. S. Roberts, and R. Thadaka, ''Prospective, Multi-Centre Trial of Mortality Following General of Spinal Anaesthesia for Hip Fracture Surgery in the Elderly,'' *British Journal of Anaesthesia* 59 (1987), 1080–1088.

Fitzgerald, J. F., P. S. Moore, and R. S. Dittus, ''The Care of Elderly Patients with Hip Fracture, *The New England Journal of Medicine,* 319 (1988), 1392–1397.

Garnick, D. W., E. R. DeLong, and H. S. Luft, ''Measuring Hospital Mortality Rates: Are 30-Day Data Enough?'' *Health Services Research* 29 (1995), 679–695.

Globerman, S., ''A Policy Analysis of Hospital Waiting Lists, *Journal of Policy Analysis and Management* 10 (1991), 247–262.

Goldman, D., ''Managed Care as a Public Cost-Containment Mechanism, *RAND Journal of Economics* 26 (1995), 277–295.

Hadley, J., E. P. Steinberg, and J. Feder, ''Comparison of Uninsured and Privately Insured Hospital Patients, *The Journal of the American Medical Association* 265 (1991), 374–379.

Hamilton, B. H., V. H. Hamilton, and N. E. Mayo, ''What Are the Costs of Queuing for Hip Fracture Surgery in Canada?'' *Journal of Health Economics* 15 (1996), 161–185.

Heckman, J., and B. Singer, ''A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data, *Econometrica* 52 (1984), 271–320.

Hoerer, D., G. Volpin, and H. Stein, ''Results of Early and Delayed Surgical Fixation of Hip Fractures in the Elderly: A Comparative Retrospective Study, *Bulletin Hospital for Joint Diseases* 53 (1993), 29–33.

Kalbfleisch, J. D., and R. L. Prentice, *The Statistical Analysis of Failure Time Data* (New York: John Wiley and Sons, 1980).

Katz, S. J., H. F. Mizgala, and H. G. Welch, ''British Columbia Sends Patients to Seattle for Coronary Artery Surgery,'' *The Journal of the American Medical Association* 266 (1991), 1108–1111.

Kenzora, J. E., R. E. McCarthy, J. D. Lowell, and C. B. Sledge, ''Hip Fracture Mortality,'' *Clinical Orthopaedics and Related Research* 186 (1984), 45–55.

Kiefer, N. M., ''Economic Duration Data and Hazard Functions,'' *Journal of Economic Literature* 26 (1988), 646–679.

Lancaster, T., *The Econometric Analysis of Transition Data* (Cambridge: Cambridge University Press, 1990).

Lindsay, B. G., ''The Geometry of Mixture Likelihoods: A General Theory,'' *Annals of Statistics* 11 (1983), 86–94.

Papadeimitropoulos, E. A., P. C. Coyte, R. G. Josse, and C. E. Greenwood, ''Current and Projected Rates of Hip Fracture in Canada, *Canadian Medical Association Journal* 157 (1997), 1357–1363.

Pickles, A., and R. Crouchley, ''A Comparison of Frailty Models for Multivariate Survival Data,'' *Statistics in Medicine* 14 (1995), 1447–1461.

Romano, P. S., L. L. Roos, and J. G. Jollis, ''Adapting a Clinical Comorbidity Index for Use with ICD-9-CM Administrative Data: Differing Perspectives,'' *Journal of Clinical Epidemiology* 46 (1993), 1075–1079.

Roos, L. L., R. K. Walld, P. S. Romano, and S. Roberecki, ''Short-Term Mortality after Repair of Hip Fracture: Do Manitoba Elderly Do Worse?'' *Medical Care* 34 (1996), 310–326.

Sabiston, D. C., *Textbook of Surgery* (Philadelphia: W. B. Saunders Company, 1991).

Seagroatt, V., and M. Goldacre, "Measures of Early Postoperative Mortality: Beyond Hospital Fatality Rates," *British Medical Journal* 309 (1994), 361–365.

U.S. Government Accounting Office, *Canadian Health Insurance: Lessons for the United States,* GAO/HRD-91-90 (1991).

Wenneker, M. B., J. S. Weissman, and A. M. Epstein, "The Association of Payer with Utilization of Cardiac Procedures in Massachusetts," *The Journal of the American Medical Association* 264 (1990), 1255–1260.

# APPENDIX A

## Data Sources

The data come from the Quebec MED-ÉCHO database and information on patients in Massachusetts drawn from the U.S. Nationwide Inpatient Sample (NIS) database.[21] All acute-care hospitals in these regions report details of each patient discharge to their respective provincial/state health authorities. Information on patients from all acute-care hospitals in Quebec is included in the MED-ÉCHO database, while the NIS contains all discharges from a random sample of 45 hospitals in Massachusetts. All patients admitted to acute-care hospitals with a primary diagnosis of hip fracture (ICD-9 codes 820.0–820.9, Fracture of neck of femur) who were discharged between calendar years 1990 and 1992 in Massachusetts and between April 1990 and March 1993 in Quebec are included in the sample.

Patients who fractured their hip due to multiple trauma (such as a motor vehicle accident) are more likely to receive expedited surgery (that is, hip fracture is not the primary health problem). Because these cases are typically handled in a special manner, we exclude them from the analysis. Thus, patients who were documented as also having had head trauma (ICD-9 codes 800–801, 803–804, or 850–854), or chest or abdominal trauma (ICD-9 codes 860–869), were excluded. In addition, only patients admitted from home or the emergency room were included in the analysis, since total wait times for transfers are not available. The restrictions yielded a sample size of 12,016 patients in Quebec and 8,979 patients in Massachusetts.

One may question the generalizability of our results to the remainder of the United States and Canada. We chose to analyze Massachusetts and Quebec for reasons of data availability and comparability. We had access to only Quebec hospital data in Canada. In addition, only four states in the NIS (California, Florida, Massachusetts, and New Jersey) had diagnosis and procedure-code data that was comparable in detail to Quebec's. We focused on Massachusetts, because it was closest in geographic proximity to Quebec.

Nevertheless, we were able to compare some basic summary statistics for hip-fracture patients in other U.S. and Canadian states and provinces as well.[22] Inpatient mortality and postsurgery length of stay ranged from a low of 2.6% and 8.4 days, respectively, in California, to a high of 5.4% and 19 days in New Jersey, with Massachusetts lying in-between (4.4% and 13.1 days). Surgery delays ranged from a low of 2.1 days in California and Wisconsin, versus a high of 3.1 days in New Jersey; Massachusetts patients faced a delay of 2.6 days on average in 1991.

Relative to the United States, Ontario, British Columbia, and Alberta reported a higher inpatient mortality rate of 6.7% and a longer total length of stay of 19.9 days in 1993–1994. Given that length of stay has been gradually falling in Canada, these figures are comparable to the 7.8% inpatient mortality rate and 28-day postsurgery length of stay identified for 1990–1993 Quebec patients in our sample. Thus, while neither Massachusetts nor Quebec is exactly identical to other states or provinces, these two regions tend to reflect their prospective countries' patterns of care and outcomes for hip-fracture patients.

[21] The NIS is part of the Agency for Health Care Policy and Research (AHCPR) Healthcare Cost and Utilization Project, 1988–1994 (HCUP-3).

[22] Data for the U.S. states was drawn from a 20% sample of all patients in the 1991 NIS who were treated in states with a substantial number of hip fracture patients (California, Florida, New Jersey, Pennsylvania, and Wisconsin). Summary statistics from Canada come from a published study of hip-fracture patients treated in Ontario, British Columbia, and Alberta in fiscal year 1993–1994 (Papadeimitropoulos et al. (1997)).

TABLE A1.—SUMMARY STATISTICS, OVERALL AND BY EXIT DESTINATION, BY COUNTRY

| | Canada | | | United States | | |
|---|---|---|---|---|---|---|
| | | Discharge Type | | | Discharge Type | |
| Variable | All | Live | Dead | All | Live | Dead |
| Undergoes Surgery | 0.92 | 1 | 1 | 0.91 | 1 | 1 |
| $t$ (mean) | 28.0 | 27.5 | 34.4 | 13.9 | 13.6 | 21.8 |
| $t$ (median) | 17 | 17 | 16 | 10 | 10 | 13 |
| $t_w$ (mean) | 5.2 | 3.0 | 4.4 | 3.5 | 2.6 | 3.1 |
| $t_w$ (median) | 2 | 2 | 2 | 2 | 2 | 2 |
| Age | 75.7 | 75.4 | 83.0 | 78.6 | 78.7 | 83.0 |
| Male | 0.28 | 0.26 | 0.35 | 0.24 | 0.23 | 0.38 |
| Pertrochanteric fracture | 0.50 | 0.49 | 0.57 | 0.52 | 0.52 | 0.54 |
| Other fracture | 0.18 | 0.18 | 0.17 | 0.12 | 0.12 | 0.12 |
| Dementia | 0.07 | 0.06 | 0.14 | 0.10 | 0.10 | 0.08 |
| Pulmonary disease | 0.11 | 0.10 | 0.20 | 0.14 | 0.13 | 0.21 |
| Diabetes | 0.09 | 0.09 | 0.12 | 0.11 | 0.11 | 0.10 |
| Liver disease | 0.01 | 0.005 | 0.02 | 0.01 | 0.003 | 0.02 |
| Charlson index | 0.54 | 0.47 | 1.30 | 0.63 | 0.60 | 1.01 |
| Teaching hospital[1] | 0.49 | | | 0.36 | | |
| Hospital surgical volume[1] (# of surgeries per year) | 54.2 | | | 63.4 | | |
| $N$ | 12016 | | | 8979 | | |

[1] Averaged over hospitals.