LICOS
Centre for Transition Economics

LICOS Discussion Papers

Discussion Paper 140/2004

Are Patients in the Transition World Paying Unofficially to Stay Longer in Hospital? Some Evidence from Kazakhstan

Robin Thompson
Ana Xavier

Katholieke Universiteit Leuven

LICOS Centre for Transition Economics
Huis De Dorlodot
Deberiotstraat 34
B-3000 Leuven
BELGIUM
TEL:+32-(0)16 32 65 98
FAX:+32-(0)16 32 65 99
http://www.econ.kuleuven.ac.be/licos

Are patients in the transition world paying unofficially to stay longer in hospital? Some evidence from Kazakhstan.[±]

Robin Thompson
Centre for Health Economics,
University of York,
York YO10 5DD, UK

Ana Xavier*
LICOS Centre for Transition Economics,
Katholieke Universiteit Leuven,
3000 Leuven, Belgium

January 2004

[±] Funding from the Department for International Development (research grant R7433), UK, the Centre for Health Economics - CHE, York, UK and the Fundação para a Ciência e Tecnologia, Lisbon, Portugal is acknowledged. The views expressed are not necessarily those of the funding bodies. We would like to thank J. Konings, A. Jones, D. Dawson and T. Ensor for most helpful comments. We thank all the participants in seminars at York and KULeuven and those at the World Congress in Health Economics, San Francisco, 2003 for their useful comments. The usual disclaimers hold.

*Corresponding author: *LICOS Centre for Transition Economics, Katholieke Universiteit Leuven, Deberiotstraat 34, 3000 Leuven, Belgium. Tel: +32 16 32 65 92. Fax: +32 16 32 65 99. Email: ana.xavier@econ.kuleuven.ac.be*

Abstract

In most of the countries in transition from a planned to a market economy (Former Soviet Union (FSU) and Central and Eastern Europe (CEE)) patients are routinely asked to pay unofficially for the medicines and other supplies that ought to be free as an attempt to improve service quality received in chronically underfunded state facilities. To empirically test whether, as surveys and anedoctal reports suggest, patients are paying to stay longer in hospital, perceived as resulting in better care (e.g. more professional attention), we construct a unique dataset on hospital length of stay, severity, unofficial payments and socio-economic characteristics (age, gender, occupation and income) from a survey on 1508 trauma and surgical patients discharged from Almaty City (the former capital of Kazakhstan) three main hospitals between 1999 and 2000. We use parametric and semi-parametric continuous time and discrete time proportional hazard and accelerated failure time models (Exponential, Gompertz, Weibull, Cox, Cloglog, LogNormal and Generalised Gamma) so as to understand whether and in which way unofficial payment is associated with discharge. We control for heteroskedasticity and conduct various diagnostic tests as well as conducting various robustness checks (unobserved heterogeneity and payment endogeneity). Results, consistent across models, suggest that patients are paying unofficially to stay longer in hospital i.e. payment decreases the hazard of leaving hospital. The hazard of being discharge is lower the higher is patient's severity and age and for women, while it is higher for private workers and housewives. Some policy implications are provided.

1 Introduction

In many emerging market economies of the former Soviet Union (FSU) and Central and Eastern Europe (CEE) patients are routinely asked and expected to pay *unofficially* for their hospital treatment, including the required medicines and medical supplies (World Bank, 1998a, 2000a). These payments have been defined as "payments to individuals or institutions in cash or in kind made outside official payment channels for services that are covered (*without direct charge*) by the public health care system" (Lewis, 2000). Often illegal, they are widespread in health care sectors in the region: patient surveys undertaken in Czech Republic, Bulgaria, Poland, Turkmenistan and Tajikistan found that respectively 27, 43, 46, 50 and 70% of the respondents paid or expected to pay for services that were officially free (Masopust, 1989; Delcheva *et al.*, 1997; Ladbury, 1997; Chawla *et al.*, 1998). The purchasing of drugs is a common source of unofficial expenditure, whilst inpatient hospital care constitutes its most costly item (Lewis, 2000, 2002).[1]

Despite their wide prevalence, little empirical analysis has so far investigated the use and implications of unofficial or informal payments. Much of the existing literature has focused on differing types of unofficial payments and the contribution these payments make to total health care spending (*e.g.* Thompson and Witter, 2000). Many of these studies are qualitative by nature and focus on small-scale interviews with patients and physicians. According to Lewis (2002), "a greater understanding is thus important if abuse of the system is to be addressed and resolved". Note moreover that, albeit the increase in the disease burden (mortality and morbidity) that paralleled the transition process from a centrally planned to a market economy, health issues concerning the population of the FSU and CEE have been subject to little research and publishing interest, especially when compared to other phenomena that characterised the transition process (*e.g.* privatisation). Indeed, the general knowledge regarding health and health care in the region held by those outside it is rather limited. We contribute to filling that gap.

Unofficial payments are rooted in systems of bargaining and connections inherited from the communist system (Smith, 1973). The rigid nature of health care provision led patients to search for mechanisms to obtain faster and better services (*e.g.* more doctor's attention) than the basic state services (Gaal, 1999a,b; Kornai, 2000). These payments also increased the reward to the medical activity, highly demanding but little recognised. Whilst transition and economic liberalisation may lead to cultural changes, it may take a fair amount of time to shake off old practices.

The recent widespread existence of unofficial payments for health care is, however, also closely related to the transition process. This led to an initial output and employment decline from which several countries have not yet fully recovered and implied the general restructuring and closure of enterprises and thus increased unemployment. This resulted in a decline in tax revenue and subsequent reductions in government health care sector funding. Government failure in addressing the scope and scale of service provision (downsizing services and reducing staff) in the context of resource constraints led to a gap in state resources necessary to fund the existing level of provision. Chronic shortages and inadequate equipment due to the tighter budget constraints resulted in patients and relatives being routinely asked to cover the shortfall in health care funding by paying, through unofficial channels, for medicines and other supplies required for their treatment. The low health workers wages compared to other state and private sector professions, the common long delays in wage payment (*e.g.* Lithuania, Ukraine and Russia) and the (practically) non-existent private sector, which could provide extra income, contributed to the increase in unofficial payments (Healy and McKee, 1997). These double the wages of physicians in Poland (Chawla *et al.*, 1998) and increase five times the salaries of doctors in Albania (Healy and McKee, 1997). They represent 62% of physicians' net income in Hungary (Kornai, 2000) and 18.5% of their monthly salary in Estonia (Barr, 1996). Minimal sanctions, weak monitoring and widespread

[1] See Thompson and Witter (2000) and Ensor (2000) for typologies of these payments.

corruption[2] fuel unofficial payments. Patients' lack of information and non-reporting and physicians' lack of accountability to a higher authority help maintaining the system.

Unofficial payments can be viewed as an attempt (an economically rational response) to improve service quality received in chronically under-funded state facilities (Thompson and Witter, 2000; Shahriari *et al.*, 2001). Improvements may include more effective medicines than those offered by the state, minimally invasive technologies rather than conventional surgery or simply motivating more "effort" by the physician (extra attention), exerting some kind of control over the latter and adapting treatment to patients' convenience: *e.g.* a better bed, a chosen doctor or a shorter wait (Field, 1989; Gaal, 1999b; Thompson and Witter, 2000; Kornai 2000; Shahriari *et al.*, 2001).

These payments play a role in sustaining health care systems in countries where, despite government efforts, public revenues generated officially have been limited (World Bank, 1998a, 2000a,b). They are a significant slice of total health care expenditure (*e.g.* 25-30% of the 1996 Kazak State budget considering medicines alone) and a financial supplement for health workers (Ensor and Savelyeva, 1998; Kornai, 2000). However, unofficial payments are likely to be inequitable, as patients' access to services or their quality depends on their ability to pay especially if the amount paid is unadjusted to patient's income (Bognar *et al.*, 2000). Indeed, they represent 3-14% of a patient's average monthly income in Bulgaria (83% if a surgical procedure is involved), while in Kazakhstan inpatient medicines may cost US$50 per patient (Delcheva *et al.*, 1997; Ensor and Savelyeva, 1998). In Russia "lack of money" was cited as the main reason for the inability to obtain medicines, twice as often in 1996 as in 1994 (Liu *et al.*, 1998). Some people may be delaying care and avoiding the health sector all together (Lewis, 2000, 2002). These payments are also deemed inefficient if channelled to individuals not to the system thus undermining investment.

---

[2] In the context of another paper one interviewed physician stated "(…) *If you want to understand informal payments (in health care) you can examine any other state structure, because our society is totally corrupt from top to bottom. Here (in Kazakhstan), the most common (example of this practice) is the road police. (…) road police always take money for anything. If I am stopped I'll give 500 KZT because the fine is 1000 KZT. The same is the case in health care. Try to introduce a law (to tackle these payments) and our masters of art will move round the law. The most surprising thing is that people don't mind; people seem to like it. It's very convenient when you have money and you can buy and solve*

In this paper, we analyse in an empirical way[3] whether unofficial payments influence hospital length of stay (LOS), which we consider a proxy for the quality of care received in the context of the transition world and certainly in Kazakhstan.[4] The investigation is motivated by the general perception that, on one hand, the quality of state health care provision is seen as poor and therefore some patients are willing to pay unofficially in an attempt to improve care quality and, on the other hand, state salaried physicians, poorly paid and unmotivated but with better information and decision power over the use of resources and patients' discharge, adjust the quality of care to the level of unofficial payment paid by patients – *e.g.* they may keep patients in hospital or discharge them early (a significant power due to the lack of follow-up service provision outside hospital) depending on the amount paid. They thus offer differing levels of service quality (*e.g.* different LOS in our paper) to paying and non-paying individuals, doing so with the knowledge that corruption is largely ignored by the state. Indeed, given harder budget constraints in the health care sector and the apathetic attitude of government towards corruption, or the inability to fight it, in some of the countries of the FSU and CEE state salaried physicians may have adopted patterns of market behaviour within state hospitals creating an unofficial market for health care.[5] This may be comparable to health markets in the Western world where price has been shown to affect the demand of services (e.g. McGuire, 2000 and Dranove and Satterthwaite, 2000). Whitney *et al.* 1997, Brennan and Spencer 2002, show that in the US and Australia respectively high priced dentists provided higher quality services such as the use of rubber dams and a shorter wait in the dental office after a patient's arrival for an appointment and Hellinger (1996) reviews a group of studies that show that fee-for-service schemes lead to higher utilisation namely the performance of more diagnostic tests.

---

anything." These highlight the entrenched nature of unofficial payments and the attempts to work around legislation as well as showing that patients and health workers are pragmatic decision-makers.

[3] Previous work by Thompson and Xavier (2002) have formalised the demand and supply sides of this unofficial market for health care quality that takes place within state care facilities.

[4] While in the USA and many EU countries post-hospital follow-up care is increasingly important as the length of hospital stay is reduced, in Kazakhstan the former is weak or non-existent and anecdotal reports suggest that patients are willing to pay to stay in hospital for reassurance.

[5] A typical "unofficial contract" is verbal and consists of few indirect words. The physician hints the patient that she has to pay for something while the patient may say "Please do everything and I will thank you". Interviews with physicians revealed that they may tell patients they need to perform various diagnostic tests, that the hospital does not have the

An important issue is whether payments are made for an enhanced level of care or entitled services (with non-paying patients receiving a sub-basic level of care). Patients may be asked to pay for medicines that hospitals do not have or are available but with a delay. Or a corrupt health worker may ask a patient for a payment to ensure access to a basic level of service implying that payment is linked to higher quality care. Information asymmetry and endemic unofficial payments place patients in a vulnerable position as the physician can abuse his knowledge of the scope and scale of care to obtain payments. Whatever the case, if the patient perceives that no payment leads to a sub-desirable care, she may be willing to pay.

This paper contributes to the existing literature in various ways. First, it uses econometric tools to analyse unofficial payments and explore whether payment influences hospital LOS (a potential proxy for the quality of care received certainly in the context of Kazakhstan) and by how much. Applying survival models to survey data on discharged acute hospital (surgery and trauma) patients in Kazakhstan, including information on patients' payment activity, social-economic characteristics and hospital LOS, we test the hypothesis that larger unofficial payments are associated with longer LOS which we see as increased quality in the specific context of Kazakhstan. Previous studies were limited to answer the "whom, how much, when and to whom", estimating spending through primary surveys. There are few if published studies in English, which formally attempt to test physician behaviour within an unofficial payments context despite the number of anecdotal reports (*e.g.* case studies in Kazak newspapers). Second, we use unique detailed data on discharged hospital patients that are generally very difficult to obtain. The data are gathered for those who had a hospital intervention that was officially free of charge allowing us to identify clearly the amount paid unofficially. Many previous studies could not distinguish between official and unofficial payments thus providing only a rough idea of what the latter might be. Finally, the work contributes to broadening the knowledge regarding health and health care in the region.

---

required supplies and beds are full or that patients have to wait a long time in the AD, all these not necessarily true. Physicians may also prescribe unnecessary medicines or an operation.

The paper is organised as follows: Section 2 describes the Kazak health care system. Section 3 describes the data and the variables used in the study while section 4 presents the duration models used in the empirical exploration. Section 5 presents the results and section 6 concludes.

2 The health care system in Kazakhstan

Before independence in 1991, the Ministry of Health in Kazakhstan administered policy made in Moscow through a centrally organised hierarchical structure, from the republic level to the *oblast* or city administration, then to the subordinate rayon level. The Kazakhstan health care system featured most of the usual characteristics of a Soviet health care system (see Ryan, 1978 for a detailed description of the organisation of Soviet health care): services were, in principle, accessible and mostly free to everyone; funding was based on capacity rather than activity; over emphasis was given to specialist training and there was a dependence on hospitalisation, with long lengths of stay; and incentives focused on penalties for failure rather than incentives for success (Ensor and Rittmann, 1997). The weaknesses of the Soviet system have been exacerbated since independence by the declining health sector spending, a product of deep economic recession. National income halved between 1991 and 1995, government revenue fell by more than 70% and real public health care spending declined dramatically (World Bank, 1998b; McKee *et al.*, 2002). Public health spending as a percentage of the GDP was lower throughout the 1990s as compared to the Soviet times. The acute funding crisis and over-emphasis on inpatient care resulted in resources being extremely thinly spread. From 1999 the country starts to recover modestly (EBRD, 2002).

Kazakhstan began the 1990s with a government funded, tax-based, health care system. A mandatory health insurance (MHI) system (based on payroll taxes earmarked for a health fund) was established in 1996. With the creation of the fund the state budget allocation to health decreased (McKee *et al.*, 2002). The MHI system was dissolved in 1998 largely due to: enterprises being unable to pay contributions to the insurance fund; a large informal workforce; the inability of regional administrations to cover the socially unprotected population (*e.g.* the growing unemployed) and a collapse in the confidence in the fund with allegations of corruption and misappropriation of reserve funds. In 1998 public spending (state budget and MHI contributions) as a percentage of the GDP was at a whole time low (1.9%) as the creation of a new source of funding did not compensate for the reduction in state funding (McKee *et al.*, 2002). Currently, funding comes from two main sources (similar to pre-insurance): the government budget and out-of-pocket payments both official (formalised by a 1999 government decree) and unofficial (European Observatory, 1999). A 1994 survey of 5000 households in South Kazakhstan found that informal payments were common for both outpatient and inpatient care. For inpatient care, payment was made to providers 11% of the times and 12% to surgeons, 25-42% of those hospitalised provided their own bedding, clean laundry and food and 57% provided their own medicines (Sari *et al.,* 2000).

Note, though, that despite the starting of the recovery the ability of a significant part of the population to pay for care is limited: a 1996 survey found that over a third of the population lived below a “subsistence minimum” living standard (World Bank, 1998b). Moreover, while entitlement to comprehensive health care was a feature of the Soviet system, in recent years entitlement benefits have become confusing. This has partly been the result of the insurance experiment where services were separated into two “packages”: basic (provided by insurance) and guaranteed (paid for by the state). Confusion is enhanced by chronic under-funding shortages and health sector corruption. In principle primary health care consultations are free, although medicines are not free for the non-exempt (the exempt include pregnant women and servicemen). Yet, even the exemption system does not function well and many individuals pay for medicines that are free. Hospital entitlement is particularly confusing and whether a patient pays depends on an illness being acute/not acute, resource availability and corruption. For example, individuals requiring elective surgery are increasingly required to pay whereas emergency patients are, in principle, exempt from payment. However, as results show, in reality a vast majority of patients pay for hospital care.

The health care system is dominated by hospital care and the length of stay in hospital appears to be important. While in countries like the UK post-hospital follow-up care is increasingly important as the length of hospital stay is reduced, in Kazakhstan the former is weak or non-existent and anecdotal reports suggest that patients are willing to pay to stay in hospital for reassurance.

3 Data

The data used in this analysis come from a survey of randomly selected 1508 discharged surgical and trauma inpatient patients treated in three hospitals in Almaty City, Kazakhstan, in 1999 (300 from hospital 1, 603 from hospital 2 and 605 from hospital 3). These are three of the main four acute hospitals in Almaty. The survey was conducted in January 2002 with a maximum of nine months elapsing between discharge and interview. Given the sensitivity of the survey (unofficial payments are part of an unofficial market) patients were surveyed in their homes. The questionnaire was conducted in Russian with the help of the staff from the School of Public Health, Kazakhstan State Medical University. Each of the patients included in the analysis is identified by an ICD10 code. Thirty-seven codes were included in the survey representing the most common surgical and trauma conditions treated in each of the hospitals. They were also chosen because individuals suffering from one of these codes were entitled to free care and thus all that was paid in hospital constituted an unofficial payment. The total population of each of the main ICD10 codes over a 12-month period was considered and 95% of those approached agreed to be surveyed. The ICD10 codes were aggregated into four resource groups in increasing degree of severity (RG1-4) based on information on resource use provided by the Almaty City Health Administration.[6] Each patient was coded in only one of the RGs. Patients were surveyed about their experience in hospital and related expenditure. They were asked if they, or their relatives on their behalf, had paid

[6] RGs are like mini DRGs and they reflect severity. RG1 – uncomplicated acute appendicitis (AC) and hernias; RG2 – AC with peritoneal abcess, more complicated hernias, ulcer disease, incised wound, suppurative diseases of skin, fractures; RG3 – AC with peritonitis, more complicated hernias, perforated ulcer, acute cholecystitis, penetrated incised wounds, hemorrhage, acute pancreatitis, other fractures; RG4 – acute ulcer disease, acute pancreatitis with complications, mechanical jaundice, hard combined trauma, intestinal fistula.

(monetarily or in kind) and the amount paid in the admission department (AD), in the surgical/trauma ward, for medicines and diagnostic tests.[7] Patients also stated amongst other things the number of nights spent in hospital. We obtained information on patients' socio-economic status: age, gender, education, occupation, exemption status and household expenditure (food, utilities, clothes, cigarettes, alcohol, cars, education, health care and drugs, celebrations and support to relatives) which we used as a proxy for household income. We also gathered information on the referral type (self-referral, polyclinic doctor or specialist, hospital specialist) and on whether the patient had surgery.

3.1 Length of hospital stay (LOS)

The empirical analysis and in the absence of a single correct, complete and tangible quality indicator (Campbell et al. 2000), focuses on length of stay (LOS) as an indicator of health care quality (process measure of quality) in that variations in LOS may point to differences in the quality of health care provision. However, we may need to distinguish between the developed world and that of transition. In OECD countries, Barnham and Kutzin (1993) argue, longer stays do not necessarily contribute to higher-quality care (although patients may not necessarily perceive it to be so): LOS for most conditions has decline during the last thirty years in most OECD countries while the health of the population has not. Improvements in the technical quality of hospital care and most importantly a much wider availability of community care and local facilities as follow-up care have made this possible[8] although concerns are sometimes raised about early discharge, post-surgical complications, and hospital readmission. In the transition world especially in the FSU the situation is quite different. Good health facilities are limited in number and often located in cities far from an important part of the population. Post-hospital follow-up is poor or non-existent and transport to hospital is limited and costly, especially from remote areas. Quality of care has

[7] We find hardly any variation in these latter two categories *i.e.* everybody appears to be paying for drugs and tests.
[8] In several OECD countries LOS has been used as a proxy for resource use and technical inefficiency.

decreased with the transition process and the consequent economic crisis. In this context, a longer stay increases patients' reassurance and decreases the probability of post-treatment complications and readmission, as doctors monitor patients for longer. Hence, a longer LOS may be perceived as better quality by patients in the transition world.

Table 1 shows LOS, by hospital and RG, whose mean is approximately 14 days. There are large differences between hospitals 1 and 2 (surgical) and hospital 3 (trauma). In hospitals 1 and 2 LOS is under 10 days while in hospital 3 LOS is over 20 days. There are also differences across RGs. The variable was standardised by ICD10 code. A negative sign is expected for the coefficient estimate associated with payment if a longer LOS proxies a perceived higher care quality.

3.2 Independent Variables

As we focus on the relationship between LOS and the unofficial payment made we gathered information on whether, how much and where patients paid. The general idea developed through the interviewing process was that payment negotiation takes place as soon as the patient arrives to hospital in the AD and before treatment takes place with patients agreeing to a certain amount for a certain quality level. However, although negotiation and agreement take place in the AD and before treatment, some patients do not pay all at once in the AD (*e.g.* because they cannot afford or have the money ready) and some pay after admission takes place while in the ward. Hence, given the information gathered with the questionnaire, we consider two unofficial payments variables: 1) Payment1 the amount of unofficial payment made by the individual in the AD and before treatment takes place; and 2) Payment2, the amount of unofficial payment made after admission when already in the ward. Both are in their logs due to their skewed distribution. We also consider payment-hospital interactions.[9]

In order to isolate the association between LOS and payment it is important to control for other factors. Martin and Smith (1996) argue that LOS is related to patient characteristics. Studies typically

[9] Note that as payment is given in units of Kazakhstani Tenge and as typically used in such analysis we have recoded the zero payments to 3 KZT when computing the log so as not to miss these observations (the idea in this context being that, given the size of payments, 3KT or 0KZT are pretty much the same thing).

find that age and severity (RG status) are important determinants of LOS (Godfarb *et al*., 1983; Cairns and Munroe, 1992) and those with a lower socio-economic status have longer LOS (Epstein *et al.*, 1990). Some physicians stated that they look at the patient's clothes, behaviour, how he/she speaks and where he/she works to establish payment ability and willingness to pay. Studies of care utilisation typically control for patient characteristics such as gender because behaviour appears to differ between man and women (Hellinger, 1996). Hence, we use age, gender, RGs, occupation, income and exemption to account for socio-economic differences or to proxy for severity.

The importance of hospital characteristics and organisational factors - discharge system, unplanned admissions, physician workload and the way by which inpatient services are financed - in determining LOS has also been established in some studies (*e.g.* Cannoodt and Knickman, 1984; Burns and Wholey, 1991; Xiao *et al.* 1997; Westert *et al.* 1993). We do not have detailed information on providers but control for differences across hospitals (*e.g.* number of beds or doctors) using dummy variables. Note that funding mechanism – now a case-based and prospective payment using RGs (Ensor and Thompson, 1999) – and medical standards indicating a patient's LOS in hospital apply equally to all the hospitals studied reducing the need to control for hospital characteristics. Finally, we also consider payment – hospital interactions so as to explore whether payment effects are different accross hospitals.

Table 2 provides a list of the variables used in the analysis. Table 3 presents some summary statistics. About 33% of the patients paid in the AD while 22% paid when already in the ward with 6% of the patients paying in both places. On average individuals paid 2,950 KZT before receiving any clinical intervention (the highest average payment of 4,266 KZT registered in hospital 2 and the lowest 1,899 KZT in hospital 1) and 1,797 KZT while already in the ward (1,931 in hospital 1, 2,099 KZT in hospital 2 and 1,425 KZT in hospital 3, average values) although payment varies across individuals. Average monthly income was around 7,220 KZT with the smallest income equal to 1,000 KZT. About 25% and 50% of the patients had respectively a monthly income of less then 4,500 and 5,780 KZT. On average payment can be a rather a high percentage of patients' monthly income (40% and 25% when considering payment in the AD and in the ward respectively). When looking only at those who paid than payment exceeds, on average, a month of income. 50% of the patients are male and average age is 43 years. 13% are students, 18% are unemployed, 12% work in private companies, 4% are self-employed, 25% are retired and 10% are housewives. About 28% of the patients are exempted from any payment for health care. About 40% of all patients go to hospital 2 and 3 while 19% go to hospital 1. More than 50% of patients are coded as RG3, 33% fall in RG2 and around 7% in each of the other two RGs.

4 Duration analysis

Given that LOS is "*time*" data we develop the empirical exploration by focusing on duration models.[10] These models are increasingly being used in health econometrics to study a range of issues such as the impact of tax on starting and quitting smoking (Forster and Jones, 2001) and the impact of hospital volume and cost on length of stay (Hamilton and Hamilton, 1997). Moreover, detailed examination of the data on LOS shows this to be positively skewed and reaching 90 days, a rather large number of hospital bed days. Hence, duration analysis may provide a good tool to better understand the data and we analyse patients' discharge, between 1999 and 2000, looking at the hazard of leaving hospital: the probability of leaving hospital at each point in time given that a patient is still in hospital at that particular time.

As LOS is measured in days the maximum number of which is 90 days the data may be considered discrete time data. Thus, we first make use of discrete time analysis estimating a Cloglog hazard function. This model is of the type: $h_j = 1 - \exp\{-\exp[x_j\beta + D(j)]\}$, where $j=1,\ldots,t$ refers to each time period (day), x is the vector of covariates, $D(j)$ is the baseline hazard (*i.e.* without considering the influence of the explanatory variables) and h_j is the hazard of being discharge from hospital at time j. This is a proportional hazard (PH) type of model.

Given the nature of the data it may nevertheless be the case that LOS data is approximated by a continuous time distribution. We therefore establish the necessary comparisons between the discrete time function above and a set of continuous parametric and semi-parametric functions (Exponential, Gompertz, Weibull, LogNormal, Generalised Gamma and Cox)[11]. Briefly, the Exponential, the Weibull, the Gompertz and the Cox are parameterised as proportional hazards using the multiplicative or PH model whereby the regressors have a multiplicative effect on the hazard of the type: $h(t_j) = h_0(t)g(x_j)$, where $g(x_j)$ is a non-negative function of the regressors (typically an exponential function: $g(x_j)=exp(x_j\beta)$). In the Cox model the $h_0(t)$ is not specified making the Cox a semi-parametric model. In the other PH models the whole expression is specified. The Exponential assumes the hazard rate is constant over time whereas the other two assume the hazard rate is either monotonically increasing or decreasing. The Exponential and the Weibull can also be expressed in the accelerated failure-time (AFT) model as are the LogNormal and the Generalised Gamma. The AFT model expresses the natural logarithm of the survival time as a linear function of the covariates: $\ln t_j = x_j\beta + z_j$, where t is the survival time (time to discharge), x is the vector of covariates and z the error term following a certain density function. If the latter is an extreme-value function, then we obtain the Exponential and the Weibull. If it is a log normal the LogNormal is obtained. The LogNormal function allows for hazard rates that are not always increasing or decreasing. The Generalised Gamma function assumes an even more flexible expression that can be reduced to the Exponential, the Weibull or the LogNormal depending on a parameter k (if $k=1$ we have the Exponential or the Weibull, if $k=0$ we have the LogNormal). Note that whereas the interpretation of the coefficient in the PH context is straightforward (one can see which variables are associated to a higher or lower hazard ratio depending on whether they present a corresponding positive or negative coefficient), the interpretation of the results when in the AFT context requires perhaps an explanation. Indeed, it is *time to failure* that is on the left hand side of the function estimated so that a positive coefficient implies a deceleration of time, that is, an increase in the time before failure occurs. We first estimate the discrete time PH model and the

---

[10] See Jones (2000) for a brief introduction to duration models in health econometrics

continuous time PH models followed by the models in the AFT format (note that we also use the Exponential and the Weibull in the AFT format as a means of comparison).

When pursuing the duration analysis we take heteroskedasticity into account in the usual way by using the White estimator of the variance. We also consider different hospital weights (which corresponds to using the option *cluster(hospital)* in STATA) as table 1 suggests that there may be differences across hospitals. Moreover, we account for potential patient unobserved heterogeneity (uing the option *frailty(gamma|invgauss)* in STATA)[12] Finally, we account for the possible endogeneity of payment in the ward (lnpayment2). Endogeneity in our context may arise from the fact that patients, once experiencing the hospital LOS, change their preferences towards the latter and therefore the payment they make in the ward (lnpayment2).[13] Hence, we use a sort of two stage estimation whereby we, first, obtain the predicted values of the payment variable by means of an OLS regression estimating payment in the ward (lnpayment2) as a function of all the other regressors and extra variables or instruments – in our case referral type and university degree – and then use the predicted values of payment rather than the observed values in the estimation of the hazard function.[14] However, we are fully aware of the fact that instruments in a cross-section context, namely based on survey data may be limited. Thus, the initial analysis may still be "best". The results are however in line with each other.

As diagnostic tests we use the normal Wald χ_2 for the joint significance of the regressors and a χ_2 test for the proportional hazard assumption of the Cox function. Where appropriate we also conduct a test for unobserved heterogeneity. We further use a Wald χ_2 test[15] to check for the general specification of the model namely for the omission of variables (Forster and Jones, 2001).

---

[11] See *e.g.* Stata 7/8 Reference Manuals for a detailed description of these duration models.
[12] Note that in the case of the Cox function we use the *cluster(patient)* option as a proxy for heterogeneity/frailty since the frailty option required a matrix size higher than the standard one provided by STATA.
[13] The definition of payment in the AD (lnpayment1) makes it exogenous.
[14] Note that we first of all wish to look for evidence of an association between payments and quality after controlling for other variables. As such, endogeneity is not an issue.

Rejecting the null hypothesis indicates that the model is not well specified. Finally, we compared the different models using the log-likelihood and Akaike information criteria (AIC). The log-likelihood criterion chooses the model specification with the highest log likelihood (and because it is a negative value the one closest to zero), while the AIC criterion implies the computation of the following expression *-2(log-likelihood)+2(c + p + 1)*, where *c* is the number of covariates and *p* the number of ancillary parameters, and chooses the model with the lowest AIC value.

5 Results

We start by examining the plot of the non-parametric hazard function (Figure 1A), defined crudely as $h_j = \frac{d_j}{n_j}$ (and assuming the covariates are zero), where *d* stands for the number of those discharged at time *j* and *n* represents the number of those still in hospital – those still at risk of being discharged - at time *j* (where *j* represents days). At first sight the hazard function plotted in Figure 1A is inconclusive. Some of the peaks may be related to for example hospital policy concerning patient discharge or just a nuisance of the data which in figure 1A do not really control for the density of patients. Indeed, looking at the density function of patients' discharge it can be seen that many patients are discharged within ten days of hospital admission, a large majority within 60 days and only a small percentage remains in hospital for more than 60 days. We therefore proceed by plotting the smooth hazard (controlling for event density) as in Figure 1B. This is based on the h_j ratio above with which we calculate the cumulative hazard H_j and the so called hazard contributions - $\Delta\hat{H}_j(t_j) = \hat{H}_j(t_j) - \hat{H}_j(t_{j=1})$ - and to which we apply a kernel (weighting) function. The resulting smooth hazard function plotted in Figure 1B appears to be increasing up to 20 days then decreasing up to 70 days, increasing again up to 80 days and decreasing once more. Again with the density in mind, we should perhaps attribute more importance to the first lump of the function. Bearing the

[15] This test consists of computing the predicted values of the dependent variable using a linear combination of the covariates, and introducing the squared values of these into the models and testing their significance. If they are

two figures in mind, we attempted to estimate the hazard of being discharge using the Exponential (constant hazard), the Gompertz and the Weibull (allowing for an increasing hazard) as well as the LogNormal (that allows us to deal with functions that are not always increasing as may be the case) and the Generalised Gamma (that allows the function to be a more complex one). We can perhaps expect that the latter two will perform better given the plots. We also plot the cumulative hazard and survival functions (Figure 2 and Figure 3).

The results for the continuous and discrete time models are presented in 4. Coefficients rather than hazard ratios are shown. Estimations are robust and different hospital weights are used to compute the standard errors. In all tables, together with the parameter estimates, various diagnostic tests and the log likelihood are presented. All models in table 4 with the exception of the discrete time Cloglog PH model and the Cox PH model pass the Wald test for general specification (namely for the omission of variables). The latter also present the worse log-likelihood value and the Cox PH model fails the proportional hazard test. This suggests that all the other model specifications are to be preferred to the Cloglog and the Cox PH models. Moreover, the **kappa** parameter of the generalised gamma is positive and significantly different from zero suggesting that the Generalised Gamma provides a better model specification than does the LogNormal. According to the log-likelihood criterion (see table 4) the Generalised Gamma specification is indeed the most preferred model followed by the Weibull, the LogNormal, the Gompertz, the Exponential, the Cloglog and finally the Cox. Using the AIC criterion the first two and the last two models keep their ranking while the Gompertz and the Exponential go up a position and the LogNormal looses two places. Note however that results are rather similar across all models namely with respect to the payment variables, which suggests that results are robust.

We find that the **Logd** parameter in the Cloglog, the **Gamma** parameter in the Gompertz, the **P** parameter in the Weibull and the **Sigma** parameter in the LogNormal and the Generalised

---

significant the model may be badly specified (e.g. Maddala, 1992).

Gamma are positive and statistically significant indicating that the hazard ratio is on average positive and increasing.

Still looking at table 4 and with respect to **unofficial payments** and their relation with the hazard of leaving hospital, the focus of our analysis, we can see that the coefficient estimates associated with paying in the AD and paying in the ward bear a positive sign across all PH models and a negative sign when in the AFT models. In other words, payment is associated with a reduced hazard of leaving hospital or an increased time in hospital before discharge. This is in line with our hypothesis that patients may be paying unofficially to stay longer in hospital (or that those not paying are discharged earlier), or, in other words, that paying patients obtain a higher quality of care than non paying patients. Note that the coefficient estimate associated with payment in the ward (**lnpayment2**) is statistically significant across all models while paying in the AD (**lnpayment1**) is statistically significantly in the discrete time model only.

Concerning the remaining variables and based on the preferred model it can be seen that the hazard of leaving hospital is lower (time to discharge is longer) in **trauma/hospital 3** as compared to **hospital 2** and is lower in both the trauma and hospital 2 as compared to **hospital 1**. This is line with table 1 and perhaps expected if we are to believe that hospital 3 receive more complex cases. Patients in diagnostic groups **RG3** and **RG4** also face a lower hazard (*i.e.* a longer LOS in hospital) as expected if these variables proxy severity. The same applies to **age. Men** present a lower hazard or a longer LOS in hospital than women do. **Income** is related to a shorter time to discharge (higher hazard ratio) the rational perhaps being that income proxies health status: those richer are in general healthier and present less complex conditions. Or else, it may be showing that those that are richer may afford to pay for home care and thus decide to stay less long in hospital. Private workers (**privwork**) and **housewives** also present a shorter time to discharge (higher hazard ratio).

We then proceed by conducting a few robust checks on these first specifications. First, we run an extended version of the models in table 4 now considering hospital – payment interactions (table 4A). Indeed, table 1 and figure 2 and 3 suggest there may be differences across hospitals and so far we have tried to accommodate those differences using hospital dummies. However, it may make sense to account for differences in the effect of payment across hospitals. Not all models pass the Wald specification test though. The models that do (the Gompertz and the LogNormal, with the LogNormal registering the highest Loglikelihood) overall confirm our previous results (only age, male and housewife are now not significant). Interestingly, we now find a statistically significant relationship between paying in the AD (**lnpayment1**) and a longer LOS in hospital. The relationship between unofficial payment in the AD and hospital stay is the strongest for hospital 3 (trauma) and the weakest for hospital 2. The relationship between unofficial payment in the ward **(lnpayment2**) and hospital stay is the strongest in hospital 1 followed closely by hospital 3 (trauma) and finally by hospital 2. We also run all the models by hospital. The results concerning the payment variables are summarised in table 4B. Again we find that unofficial payment is associated with a longer stay in hospital before discharge or a lower hazard of leaving hospital. Results confirm that the relationship between unofficial payment in the ward and hospital stay is indeed stronger in hospital 1 than in hospital 3 (trauma), which is stronger than that in hospital 2. We also find a significant relationship between paying in the AD (**lnpayment1**) and a longer LOS (lower hazard of discharge) in hospital 3 in a very similar fashion to that between paying in the ward and hospital discharge. The observed differences between the hospitals may reflect different policies followed by hospital managers to cut down on corruption. If so, hospital 2 appears to be more successful in counteracting unofficial payments.

As a further robustness check we re-estimate all the models in table 4 but now taking patient level frailty or unobserved heterogeneity into account (table 5). The estimated model with unobserved patient heterogeneity will produce an estimate of the variance of the frailties (**gamma variance** for the discrete time model and **lntheta** for the continuous time models) and an associated inference test. A likelihood-ratio test (the null hypothesis being that the variance in patient heterogeneity is zero) is also computed when in the discrete time model. If the null hypothesis is true, the model reduces to the model without frailty.

Note that not all models pass the Wald test for general specification. Those that do are coincident with those where heterogeneity is accepted (the coefficient estimates for the parameter theta are significant) with the exception of the discrete time Cloglog which does not pass the Wald test. Indeed, the log-likelihood ratio test suggests that heterogeneity exists (confirmed by the coefficient estimate for the gamma variance) but the model is misspecified. These results suggest that the Cloglog, the Exponential, the Gompertz and the LogNormal perform better when taking heterogeneity into account whereas the other models performed better in their initial specification.[16] Judging by the Log-likelihood and the Akaike information criterion for the frailty models that are deemed well specified the LogNormal is preferred to the Gompertz which is preferred to the Exponential. Nevertheless, the model with the highest Log-likelihood and the lowest AIC of the two tables is still the Generalised Gamma specification in table 4. More importantly, when looking at the coefficient estimates of the well specified models we can conclude that all our previous conclusions remain (the sign, the statistical significance of the coefficient estimates and even the magnitude are very similar indeed!), that is, results are robust. We then run the models with the interactions accounting for heterogeneity. As before, only the Gompertz and the LogNormal are deemed well specified and results remain unchanged (table 5A). We also run all the models by hospital using the frailty option. Previous concluding remarks hold and are summarised in table 5B.

As a final robust check we take endogeneity into account (table 6) and re-run the models of table 4. All models except the Cloglog and the Cox are deemed well specified both in the simple specification and in that with interactions. As previously, unofficial payment is found to be positively related to LOS in hospital *i.e.* negatively associated with the hazard of being discharged.

---

[16] Note that while estimating the Generalised Gamma with frailty we encountered some convergence problems, which remain despite various attempts to change the maximum likelihood criteria. Hence we do not report some of the diagnostic tests. It was suggested to us that this is due to the fact that the generalised gamma with frailty is trying to

Interestingly, we now find a statistically significant relationship between paying in the AD and LOS in several models while the relationship between paying in the ward and LOS is now stronger (higher magnitude of the coefficient estimates). The models with interactions confirm that the relationship between unofficial payment in the ward and hospital stay is stronger for hospital 1 followed closely by hospital 3 (trauma) and finally by hospital 1. The relationship between payment in the AD and LOS is the strongest for hospital 3 (trauma) and the weakest for hospital 2. The analysis by hospital sustains these conclusions. We find that coefficient estimates associated with the ward payment are statistically significant and similar in hospital 1 and 3 while smaller in hospital 2. AD payment is statistically significant in hospital 1 and 3 and stronger in the context of the latter. Note however that the models concerning hospital 3 appear misspecified. We also re-run the models taking heterogeneity into account. As results are mostly unchanged we restrict the presentation of the models with endogeneity to the simplest specification.

6 Conclusions

In this paper we empirically investigate unofficial payments for health care in the transition world using the example of Kazakhstan. We attempt to verify if indeed patients are paying unofficially to see an improvement in the quality of care they receive, the latter measured in terms of hospital length of stay (LOS) particularly in the case of Kazakhstan and in the absence of a single correct, complete and tangible quality indicator. The motivation for the study follows claims that patients perceive state health care provision to be poor and are willing to pay unofficially in an attempt to improve the quality of care received, while physicians, underpaid and unmotivated, have more information and power over treatment and discharge which they exploit offering differing levels of service quality to paying and non-paying patients), and doing so with the knowledge that corruption is largely ignored by the state.

---

estimate too many parameters and cannot separately identify a frailty effect. In this case it should be sufficient to stick with the other (simpler) specification.

We make use a unique data set from a survey of 1508 discharged patients treated in three hospitals in Almaty City, Kazakhstan, for conditions whose treatment was officially free of charge. Given the sensitivity of the topic they were interviewed in their homes and shortly after discharge. We gathered information on their social and economic status (age, gender, occupation and income) and their experience in hospital (diagnostic group, payment and LOS). LOS was on average 14 days with hospital 1 and 2 registering on average a LOS of under 10 days and hospital 3 over 20 days. Maximum LOS was 90 days. As LOS can be seen as "*time*" data we use a set of discrete time proportional hazard and continuous time proportional hazard or accelerated time failure models: Cloglog, Exponential, Gompertz, Weibull, Cox, LogNormal and Generalised Gamma and establish the necessary comparisons using the log-likelihood criterion and the AIC. We control for potential heteroskedasticity, patient unobserved heterogeneity or frailty and the endogeneity of the payment in the ward, and conduct various diagnostic tests namely the Wald χ_2 test for the general specification of a model.

Results suggest that unofficial payments are associated with a reduced hazard of leaving hospital or an increased time in hospital (LOS) before discharge. This confirms our hypothesis that patients may be paying unofficially to stay longer in hospital (or that those not paying are discharged too early) or, in other words, that paying patients receive a higher quality of care in detriment of non-paying patients, if LOS does indeed proxies quality of care. The coefficient estimates associated with paying in the AD and paying in the ward bear a positive sign across all models PH models and a negative sign when in the AFT models. The coefficient estimate associated with ward payment is statistically significant across all models. When considering hospital-payment interactions and looking at each hospital separately the following are found to be statistically significant. The relationship between paying unofficially in the AD and hospital stay is the strongest for hospital 3 (trauma) and the weakest for hospital 2, whereas the relationship between paying unofficially in the ward and hospital stay is the strongest in hospital 1 followed closely by hospital 3 (trauma) and then by hospital 2. This perhaps reflects different attitudes and policies implemented by hospital management to combat corruption within their hospital. Results are robust across all specifications and estimation methods. We also find evidence that a higher severity proxied by RGs and age are associated with a lower hazard. Men present a lower hazard than women do, while private workers and housewives face a higher hazard ratio. Income is related to a shorter time to discharge (higher hazard ratio) the rational perhaps being that income proxies health status: those richer are in general healthier and present less complex conditions. Or, it may indicate that richer patients may afford to pay for home care thus needing to stay less long in hospital. Finally, evidence obtained with the duration analysis is consistent with that obtained in the context of a different paper using simple OLS analysis and ordered probit to investigate unofficial payments.

This paper contributes to the existing literature in various ways. First, it uses econometric tools to analyse unofficial payments and explore whether payment influences quality of care proxied by hospital LOS. We therefore test the hypothesis that larger unofficial payments are associated with longer LOS by applying a variety of survival models to survey data on discharged acute hospital patients in Kazakhstan. We also conduct various diagnostic tests and robust checks. Second, we use unique detailed data on discharged hospital patients that are generally very difficult to obtain. The data are gathered for those who had a hospital intervention that was officially free of charge allowing us to identify clearly the amount paid unofficially. Many previous studies could not distinguish between official and unofficial payments thus providing only a rough idea of what the latter might be. Note that much of the existing literature focused on differing types of unofficial payments and the contribution these payments make to total health care spending answering to the "whom, how much, when and to whom" and estimating spending through primary surveys. There are few if published studies in English, which formally attempt to test physician behaviour within an unofficial payments context despite the number of anecdotal reports (*e.g.* case studies in Kazak newspapers). Many of these studies are qualitative by nature and focus on small-scale interviews with patients and physicians. Little empirical analysis has so far investigated the use and implications of unofficial or informal payments. This work attempts to fill that gap and broaden the knowledge regarding health and health care in the region.

Deeply rooted unofficial or informal payments for health care will most certainly remain a reality in the short run and a considerable challenge to policy makers. Although perhaps constituting a rational response to resource shortages, they may be inequitable and inefficient and thus it may be necessary to implement a whole range of policies so as to address such difficult matter. These policies may include providing more information to patients regarding their entitlement to free/guarantee packages of health care and the existing formal charges, more monitoring and the enforcement of sanctions. Moreover, to tackle unofficial payments, insofar as they are associated with financial deficiencies, implies the need to increase state revenues by addressing tax and national health insurance evasion and, due to the still difficult economic condition, to introduce/increase formal charges (using modest levels to start with) and to couple them with free packages and exemption schemes to account for the potential inequity. As part of general health reform governments may need to adjust resource use and payments systems (*e.g.* using DRGs) in order to control costs. Ultimately, a sustained economic growth will help raise state revenue and allow for an increase in the official income of state health care workers.

References

Barnham, H. and Kutzin, J. (1993). *Public Hospitals in Developing Countries*, John Hopkins University Press, Baltimore, Maryland.

Barr, N. (1996) The ethics of Soviet medical practice: behaviour and attitudes of physicians in Soviet Estonia, *Journal of Medical Ethics* 22: 33-40.

Bognar, G., Robert, I. and Kornai, J. (2000). Gratitude payments in the Hungarian health sector, *Kozgazdasagi Szemle* 47: 293-320.

Brennan, D.S. and Spencer, A.J. (2002). Influence of patient, visit and oral health factors on dental service provision, *Journal of Public Health Dentistry* 62(3): 148-157.

Burns, L. and Wholey, D, (1991). The effects of patient, hospital and physician characteristics on length of stay and mortality, *Medical Care* 293: 251-271.

Cairns, J. and Monroe, J. (1992). Why does length of stay vary for orthopaedic surgery?, *Health Policy* 223: 297-306.

Campbell, S., Roland, M. and Buetow, S (2000). Defining quality of care, *Social Science and Medicine* 51: 1611-1625.

Cannoodt, L. and Knickman,J. (1984). The effect of hospital characteristics and organisational factors on pre- and postoperative lengths of hospital stay, *Health Services Research* 195: 561-585.

Chawla, M., Berman, P. and Kawiorska, D. (1998). Financing health services in Poland: New evidence on private expenditures, *Health Economics* 7: 337-346.

Delcheva, E. Balabanova, D. and McKee, M. (1997). Under-the-counter payments for health care: Evidence from Bulgaria, *Health Policy* 42: 89-100.

Dranove, D. and Satterthwaite, M. (2000). The industrial organization of health care markets. In Culyer, A.J. and Newhouse, J.P. (eds) *Handbook of Health Economics*, Volume 1, Chapter 20, Elsevier Science, Amsterdam

Ensor, T. (2000). The unofficial business of health care in transitional Europe*, Eurohealth* 6(2) Spring Issue

Ensor, T. and Rittmann, J. (1997) Reforming health care in the Republic of Kazakhstan, *International Journal of Health planning and Management* 12: 219-234.

Ensor, T and Savelyeva, L. (1998). Informal payments for health care in the Former Soviet Union: some evidence from Kazakstan, *Health Policy and Planning* 13(1): 41-49.

Ensor, T., Thompson, R. (1999). Rationalising rural hospital services in Kazakstan, *International Journal of Health Planning and Management* 14: 155-167.

Epstein, A., Stern, A. and Weissman, J. (1990). Do the poor cost more? A multi-hospital study of patients' socio-economic status and use of hospital resources, *New England Journal of Medicine* 32216: 1122-1128.

EBRD (2002) Transition Report 2002. European Bank for Reconstruction and Development.

European Observatory (1999). *Health Systems in Transition: Kazakhstan*, European Observatory on Health Care Systems.

Field, M. (1989). The position of the Soviet physician, *Milbank Quarterly* 66(2): 182-201.

Forster, M. and Jones, A. (2001) Starting and quitting smoking*, Journal of the Royal Statistical Society*, 164(3): 517-547.

Gaal, P. (1999a). Informal payments in the Hungarian health services, mimeo.

Gaal, P. (1999b). Under-the-table payment and health care reform in Hungary, mimeo.

Godfarb, M., Hornbrook, M. and Higgins,C. (1983). Determinants of hospital use: A cross-diagnostic analysis, *Medical Care* 211: 48-66.

Hamilton, B. and Hamilton, V. (1997). Estimating surgical volume-outcome relationships applying survival models: accounting for frailty and hospital fixed effects, *Health Economics* 6: 383-395.

Healy, J. and McKee, M. (1997). Health sector reform in Central and Eastern Europe, *Health Policy and Planning* 12 (4) 286-295.

Hellinger, F.J. (1996). The impact of financial incentives on physician behaviour in managed care plans: a review of the evidence, *Medical Care Research and Review* 53: 294-314.

Jones, A (2000) Health econometrics in Culyer, A.J. and Newhouse, J.P. (eds) *Handbook of Health Economics*, Volume 1, Chapter 6, Elsevier Science, Amsterdam.

Kornai, J. (2000). Hidden in an envelope: gratitude payments to medical doctors in Hungary. Mimeo for the Festschrift in honour of George Soros.

Ladbury, S. (1997). *Social Assessment Study: Turkmenistan.* World Bank report mimeo.

Lewis, M. (2000), *Who Is Paying for Health Care in Eastern Europe and Central Asia?* World Bank

Lewis, M. (2002) Informal Payments in Central and Eastern Europe and the Former Soviet Union: Issues, Trends, and Policy implications. Mossialos et al. (eds). Open University Press.

Liu, Y., Rao, K. and Fei, J. (1998). Economic transition and health transition: comparing China and Russia, *Health Policy* 44: 103-122.

Maddala, G.S. (1992). Introduction to Econometrics", 2nd Edition, New York: MacMillan Publishing Company

Martin, S. and Smith, P. (1996). Explaining variations in inpatient length of stay in the National Health Service, *Journal of Health Economics* 15: 279-304.

Masopust, V. (1989). Bribes in health care and patients opinions, Medline abstract, source Cesk Zdrav 37 (6-7): 299-307.

McGuire (2000). Physician Agency in Culyer, A.J. and Newhouse, J.P. (eds) *Handbook of Health Economics*, Volume 1, Chapter 9, Elsevier Science, Amsterdam.

McKee, M., Healy, J. and Falkingham, J. (2002). *Health Care in Central Asia*. McKee, M., Healy, J. and Falkingham, J. (eds), Open University Press, Buckingham.

Ryan, M. (1978). The organisation of Soviet medical care. Professional Seminar Consultants, Inc.

Sari, N., Langenbrunner, J. and Lewis, M. (2000). Affording out-of-pocket payments for health services. *Eurohealth*, 6:2: Spring Issue.

Shahriari, H., Belli, P. and Lewis, M. (2001) Institutional Issues in Informal Health Payments in Poland. World Bank Draft Report.

Smith, H. (1973). *The Russians.* Sphere books.

Stata Press (2001). *Stata 7: Reference manuals.* Stata Corporation

Stata Press (2003). *Stata 8: Reference manuals.* Stata Corporation

Thompson, R. and Witter, S. (2000). Informal payments in transition economies: implications for health sector reform, *International Journal of Health Planning and Management* 15: 169-187.

Thompson R. and Xavier, A. (2002). Unofficial payments for acute state hospital care in Transition Economies. A model of price discrimination and vertical service differentiation applied to Kazakhstan. DP 124/2002, LICOS, Katholieke Universiteit Leuven.

Westert, G., Niebor,A. and Groenewegan, P. (1993). Variation in duration of hospital stay between hospitals and between doctors within hospitals. *Social Science and Medicine* 376: 833-839.

Whitney, C.W., Milgrom, P. and Conrad, D. *et al.* (1997). The relationship between price of services, quality of care, and patient time costs for general dental practice. *Health Services Research* 31(4): 773-790.

World Bank. (1998a). Health Strategy Paper for Eastern Europe and Central Asia

World Bank. (1998b). *Kazakhstan: Living Standards During the Transition.* Report No: 17520-KZ.

World Bank (2000a). Making Transition Work for Everyone: Poverty and Inequality in Europe and Central Asia.

World Bank (2000b). “Health” chapter in Czech Republic: Public expenditure review, mimeo.

Xiao, J., Douglas, D., Lee, A. and Vemuri, S. (1997). A Delphi evaluation of the factors influencing length of stay in Australian hospitals, *International Journal of Health Planning and Management* 12: 207-218.

Tables

Table 1: Days spent in hospital by hospital and resource group (mean, (standard deviation), number of observations)

Hospital	Resource Group 1	Resource Group 2	Resource Group 3	Resource Group 4	Total
1	6.0 (3.8) 24	5.9 (4.1) 107	7.8 (5.4) 126	8.6 (4.0) 42	7.0 (4.8) 299
2	6.3 (2.5) 84	7.9 (6.4) 266	10.4 (5.3) 211	14.3 (6.2) 41	9.0 (6.0) 602
3	3.0 (0.0) 1	23.0 (25.8) 127	22.8 (19.2) 447	19.1 (18.6) 27	22.7 (20.7) 602
Total	6.2 (2.8) 109	11.3 (15.5) 500	17.1 (16.3) 784	13.3 (11.0) 110	**14.1** (15.5) 1503

Table 2: Variables used in the empirical analysis

Variable code	***Description***
Dependent	Based on the LOS in hospital, *i.e.* the number of days an individual spends in hospital
H_j	The hazard of leaving hospital in period *j* for the PH models.
Ln t_j	The log of survival time, *t*, in period *j* for the AFT models.
Independent	
Socio-economic variables	
Age	Age, in years
Male	Binary gender, male = 1
Student, unemploy, statwork, privwork, selfwork, retired, housewife	Student, unemployed, state employee, private company employee, self employed, retired, housewife (Dummy variables)
Exempt	Registered exempt = 1(Dummy variable)
Lnincome	The log of the household adjusted monthly consumption expenditure (income proxy) in local currency (KZT)
Payment variables:	
Lnpayment1 **Lnpayment2**	The log of the amount of KZT paid in the Admission Dept The log of the amount of KZT paid in the ward
Hospital specific variables	
Hospital 1, Hospital 2, Hospital 3 (Trauma)	Hospitals 1, 2 and 3 (Dummy variables)
Hosp1*lnpayment, Hosp2*lnpayment, Hosp3*lnpayment	Hospital payment interactions

Table 3: Descriptive Statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
LOS	1496	13.556	13.081	1	90
Age	1508	42.989	18.004	5	89
Male	1508	0.505	0.500	0	1
Income	1494	7219.674	5189.935	1000	53517.880
Lnincome	1494	8.720	0.546	6.908	10.888
Payment1	1452	2949.345	6145.867	0	52000
Payment2	1483	1796.129	4743.295	0	35000
LnPayment1	1452	3.397	3.528	1.099	10.859
LnPayment2	1483	2.538	3.025	1.099	10.463
Student	1508	0.133	0.340	0	1
Unemploy	1508	0.184	0.387	0	1
Privwork	1508	0.117	0.322	0	1
Selfwork	1508	0.036	0.188	0	1
Retired	1508	0.253	0.435	0	1
Houswife	1508	0.102	0.303	0	1
Exempt	1508	0.284	0.451	0	1
RG1	1508	0.072	0.259	0	1
RG2	1508	0.334	0.472	0	1
RG3	1508	0.521	0.500	0	1
RG4	1508	0.074	0.261	0	1
Hospital 1	1508	0.199	0.399	0	1
Hospital 2	1508	0.400	0.490	0	1
Hospital 3	1508	0.401	0.490	0	1

Table 4: Comparison between discrete and continuous time models (different specifications of the Hazard function). Coefficients (rather than ratios) are presented.

	Proportional Hazard Form					Accelerated Failure-Time Form			
Continuous Payment	**Discrete time PH Cloglog**	**Continuous time Exponential**	**Continuous time Gompertz**	**Continuous time Weibull**	**Continuous time Cox**	**Continuous time Exponential**	**Continuous time Weibull**	**Continuous time LogNormal**	**Continuous time Gener. Gamma**
Hazard	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Logd	0.5717a								
Trauma	-1.6460a	-0.9994a	-1.3377a	-1.7300a	-1.4483a	0.9994a	1.0561a	0.8489a	0.9739a
Hac	-0.3804a	-0.2496a	-0.2724a	-0.3894a	-0.3945a	0.2496a	0.2377a	0.2593a	0.2531a
Rg1	0.2845a	0.1222	0.1723b	0.3082a	0.2497b	-0.1222	-0.1882b	-0.0390	-0.1015
Rg3	-0.2784a	-0.2067a	-0.1941a	-0.2700a	-0.2979a	0.2067a	0.1648a	0.2725a	0.2219a
Rg4	-0.5046a	-0.3483a	-0.3901a	-0.5207a	-0.5337a	0.3483a	0.3179a	0.3933a	0.3588a
Age	-0.0093a	-0.0055b	-0.0057c	-0.0093	-0.0093b	0.0055b	0.0057c	0.0033	0.0052b
Male	-0.1905a	-0.1095b	-0.1356b	-0.1959c	-0.1767a	0.1095b	0.1196b	0.0879a	0.1051b
Lnincome	0.2073a	0.1328a	0.1619a	0.2191b	0.1930a	-0.1328a	-0.1338a	-0.1325a	-0.1327a
Lnpay1	**-0.0332a**	**-0.0230**	**-0.0312**	**-0.0349**	**-0.0281**	**0.0230**	**0.0213**	**0.0278**	**0.0238**
Lnpay2	**-0.0770a**	**-0.0518a**	**-0.0631a**	**-0.0813a**	**-0.0719a**	**0.0518a**	**0.0497a**	**0.0597a**	**0.0530a**
Student	0.1108	0.0905	0.1236	0.1200	0.1046	-0.0905	-0.0732	-0.1375	-0.0976
Unemploy	-0.0409	-0.0044	-0.0299	-0.0283	-0.0176	0.0044	0.0173	-0.0397	-0.0025
Privwork	0.1380	0.0886a	0.1106a	0.1553a	0.1445a	-0.0886a	-0.0948a	-0.0921a	-0.0869a
Selfwork	0.1061	0.0461	0.1137	0.1089	0.0282	-0.0461	-0.0665	-0.0220	-0.0397
Retired	-0.0607	-0.0545	-0.0468	-0.0440	-0.0742	0.0545	0.0269	0.1222c	0.0666
Houswife	-0.1682	-0.1026a	-0.1149a	-0.1654a	-0.1560a	0.1026a	0.1009b	0.1108b	0.1035a
Exempt	-0.1045	-0.0399	-0.1200	-0.1357	-0.0413	0.0399	0.0829	-0.0136	0.0263
_cons	-3.5494a	-2.474a	-2.7673a	-4.2543a		2.474a	2.5973a	2.3521a	2.4389a
Gamma/ P/ Sigma			0.0238a	1.6380a			1.6380a	0.6833b	0.6385a
Kappa									0.4971a
No of observations	18916	18906	18906	18906	18906	18906	18906	18906	18906
Wald test for joint signif. Of variables	chi2(18)=750.573 Prob>chi2=0.0000	chi2(17)=448.92 Prob>chi2=0.0000	chi2(17)=558.98 Prob>chi2=0.0000	chi2(17)=822.50 Prob>chi2=0.0000	chi2(17)=718.93 Prob>chi2=0.0000	chi2(17)=448.92 Prob>chi2=0.0000	chi2(17)=1312.65 Prob>chi2=0.0000	chi2(17)=765.05 Prob>chi2=0.0000	chi2(17)=989.06 Prob>chi2=0.0000
Wald test specification Ho: model well specified	Chi2(1)=119.18 Prob>chi2=0.0000	chi2(1)=1.08 Prob>chi2=0.2987	chi2(1)=1.79 Prob>chi2=0.1810	chi2(1)=1.59 Prob>chi2=0.2078	chi2(1)=13.47 Prob>chi2=0.0002	chi2(1)=1.08 Prob>chi2=0.2987	Chi2(1)=1.89 Prob>chi2=0.1688	Chi2(1)=0.41 Prob>chi2=0.5223	chi2(1)=0.81 Prob>chi2=0.3687
Test proport. Haz. assumption (global test)					chi2(17)=203.21 Prob>chi2=0.0000				
Log-likelihood	-4701.9914	-1719.8305	-1656.7566	-1476.1082	-8715.7555	-1719.8305	-1476.1082	-1478.2255	-1440.155
AIC	9441.983	3475.661	3351.513	2990.216	17467.51	3475.661	2990.216	2994.451	2920.31

Notes: a, b, and c stand for significance level of 1%, 5% and 10% respectively. Estimations are robust. Wald test for kappa=1 (Ho:kappa=1) : Chi2(1)=8.83 and Prob>chi2=0.0030.

Table 4A: Comparison between discrete and continuous time models with interaction terms. Coefficients are presented.

	Proportional Hazard Form					Accelerated Failure-Time Form			
Continuous Payment	**Discrete time PH Cloglog**	**Continuous time Exponential**	**Continuous time Gompertz**	**Continuous time Weibull**	**Continuous time Cox**	**Continuous time Exponential**	**Continuous time Weibull**	**Continuous time LogNormal**	**Continuous time Gener. gamma**
Hazard	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Logd	0.6035a								
Hosp1 - Trauma	-1.5903a	-0.9161a	-1.2234a	-1.6531a	-1.4131a	0.9161a	0.9901a	0.7406a	0.8930a
Hosp2 - Hac	-0.7247a	-0.4575a	-0.5204a	-0.7446a	-0.7245a	0.4575a	0.4460a	0.4778a	0.4610a
Rg1	0.2529b	0.1058	0.1524b	0.2750b	0.2187c	-0.1058	-0.1647b	-0.0376	-0.0920
Rg3	-0.2654a	-0.2001a	-0.1777a	-0.2554a	-0.2909a	0.2001a	0.1529a	0.2656a	0.2125a
Rg4	-0.5504a	-0.3702a	-0.4240a	-0.5694a	-0.5789a	0.3702a	0.3410a	0.4091a	0.3776a
Age	-0.0089a	-0.0052b	-0.0052	-0.0090	-0.0091b	0.0052b	0.0054c	0.0031	0.0050b
Male	-0.1772a	-0.0990b	-0.1258c	-0.1814c	-0.1624b	0.0990b	0.1086b	0.0723	0.0953b
Lnincome	0.1864a	0.1179a	0.1426a	0.1966c	0.1739b	-0.1179a	-0.1178b	-0.1181a	-0.1181a
Lnpay1	**-0.0019**	**-0.0078a**	**-0.0063c**	**-0.0011**	**-0.0016**	**0.0078a**	**0.0007**	**0.0200a**	**0.0099a**
Hacalnpay1	**0.0067**	**0.0090a**	**0.0073a**	**0.0058**	**0.0073a**	**-0.0090a**	**-0.0035**	**-0.0172a**	**-0.0104a**
Traumaalnpay1	**-0.0841a**	**-0.0493a**	**-0.0712a**	**-0.0894a**	**-0.0746a**	**0.0493a**	**0.0535a**	**0.0435a**	**0.0481a**
Lnpay2	**-0.1334a**	**-0.0809a**	**-0.0969a**	**-0.1381a**	**-0.1273a**	**0.0809a**	**0.0827a**	**0.0817a**	**0.0850a**
Hacalnpay2	**0.09311a**	**0.0538a**	**0.0641a**	**0.0971a**	**0.0904a**	**-0.0538a**	**-0.0581a**	**-0.0492a**	**-0.0528a**
Traumaalnpay2	**0.0522b**	**0.0207a**	**0.0228a**	**0.0490a**	**0.0511a**	**-0.0207a**	**-0.0294a**	**-0.0048a**	**-0.0184a**
Student	0.1521	0.1109	0.1536	0.1618	0.1400	-0.1109	-0.0969	-0.1544c	-0.1157
Unemploy	-0.0483	-0.0069	-0.0391	-0.0373	-0.0200	0.0069	0.0223	-0.0366	0.0012
Privwork	0.1228	0.0862a	0.1026a	0.1380a	0.1355a	-0.0862a	-0.0827a	-0.0971a	-0.0867a
Selfwork	0.0738	0.0277	0.0843	0.0711	0.0039	-0.0277	-0.0426	-0.0061	-0.0237
Retired	-0.0595	-0.0496	-0.0518	-0.0451	-0.0622	0.0496	0.0270	0.1135	0.0579
Houswife	-0.1249	-0.0705	-0.0799	-0.1205	-0.1118c	0.0705	0.0722	0.0779	0.0707
Exempt	-0.1698	-0.0743	-0.1749	-0.2050	-0.1022	0.0743	0.1227	0.0131	0.0623
_cons	-3.3423a	-2.3151a	-2.5896a	-4.0555a		2.3151a	2.4291a	2.2075a	2.2915a
Gamma/ P/ Sigma			0.0250a	1.6696a			1.6696a	0.6744b	0.6271a
Kappa									0.5175a
No of observations	18916	18906	18906	18906	18906	18906	18906	18906	18906
Wald test for joint signif. Of variables	chi2(22)=794.4746 Prob>chi2=0.000	chi2(21)=457.13 Prob>chi2=0.0000	chi2(21)=575.19 Prob>chi2=0.0000	chi2(21)=849.34 Prob>chi2=0.0000	chi2(21)=758.96 Prob>chi2=0.0000	Chi2(21)=457.13 Prob>chi2=0.0000	Chi2(21)=1398.86 Prob>chi2=0.0000	Chi2(21)=796.38 Prob>chi2=0.0000	Chi2(21)=1039.66 Prob>chi2=0.0000
Wald test specification Ho: model well specified	Chi2(1)=130.38 Prob>chi2=0.0000	chi2(1)=9.39 Prob>chi2=0.0022	chi2(1)=0.27 Prob>chi2=0.6018	chi2(1)=5.64 Prob>chi2=0.0175	chi2(1)=9.03 Prob>chi2=0.0027	chi2(1)=9.39 Prob>chi2=0.0022	Chi2(1)=7.27 Prob>chi2=0.0070	Chi2(1)=3.34 Prob>chi2=0.0676	Chi2(1)=9.08 Prob>chi2=0.0026
Test proport. Haz. (global test)					chi2(21)=170.23 Prob>chi2=0.0000				
Log-likelihood	-4680.0406	-1711.4291	-1642.2553	-1452.1153	-8697.2783	-1711.4291	-1452.1153	-1459.5583	-1419.4024
AIC	9398.081	3458.858	3322.511	2942.231	17430.56	3458.858	2942.231	2957.117	2878.805

Notes: a, b, and c stand for significance level of 1%, 5% and 10% respectively. Estimations are robust. Wald test for kappa=1 (Ho:kappa=1): Chi2(1)=7.88; Prob>chi2=0.0050.

Table 5: Comparison between discrete and continuous time models accounting for heterogeneity (frailty). Coefficients are shown.

	Proportional Hazard Form					Accelerated Failure-Time Form			
	Discrete time	Continuous time	Continuous time	Continuous time	Continuous time	Continuous time	Continuous time	Continuous time	Continuous time
	PH Cloglog Frailty	Exponential Frailty	Gompertz Frailty	Weibull Frailty	Cox Frailty†	Exponential Frailty	Weibull Frailty	LogNormal Frailty	Gener. gamma Frailty
Hazard	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Logd	1.1293a								
Trauma	-2.0325a	-0.9993a	-1.3377a	-2.2167a	-1.4483a	0.9993a	0.9568a	0.8488a	0.9054
Hac	-0.5499a	-0.2496a	-0.2724a	-0.6036b	-0.3945a	0.2496a	0.2605a	0.2593a	0.2662
Rg1	0.1213	0.1222	0.1723b	0.1095	0.2497a	-0.1222	-0.0473	-0.0390	-0.0662
Rg3	-0.5565a	-0.2067a	-0.1941a	-0.6064b	-0.2979a	0.2067a	0.2617a	0.2725a	0.2200a
Rg4	-0.8318a	-0.3483a	-0.3901a	-0.9323	-0.5337a	0.3483a	0.4024a	0.3933a	0.3410
Age	-0.0135a	-0.0055b	-0.0057c	-0.0140c	-0.0093a	0.0055b	0.0061a	0.0033	0.0025a
Male	-0.2223a	-0.1095b	-0.1356b	-0.2274c	-0.1767a	0.1095b	0.0981a	0.0879a	0.0723
Lnincome	0.2844a	0.1328a	0.1619a	0.3120a	0.1930a	-0.1328a	-0.1347a	-0.1324a	-0.1297a
Lnpay1	**-0.0471a**	**-0.0230**	**-0.0311**	**-0.0535**	**-0.0281a**	**0.0230**	**0.0231**	**0.0278**	**0.0233**
Lnpay2	**-0.1059a**	**-0.0518a**	**-0.0631a**	**-0.1189a**	**-0.0712a**	**0.0518a**	**0.0513a**	**0.0597a**	**0.0546a**
Student	0.1539	0.0905	0.1236	0.1794c	0.1046	-0.0905	-0.0774	-0.1374	-0.1153
Unemploy	-0.0036	-0.0044	-0.0299	0.02978	-0.0176	0.0044	-0.0129	-0.0397	0.0239
Privwork	0.1427	0.0886a	0.1106a	0.1675a	0.1445	-0.0886a	-0.0723a	-0.0921a	-0.0455
Selfwork	0.0437	0.0461	0.1137	0.0248	0.0282	-0.0461	-0.0107	-0.0219	0.0140
Retired	-0.2163	-0.0545	-0.0469	-0.2349	-0.0742	0.0545	0.1014	0.1222c	0.0835
Houswife	-0.2199	-0.1026a	-0.1149a	-0.2169a	-0.1560	0.1026a	0.0936c	0.1108b	0.1203
Exempt	0.0927	-0.0399	-0.1200	0.0842	-0.0413	0.0399	-0.0363	-0.0137	0.0811
_cons	-4.2933a	-2.4739a	-2.767a	-5.4947a		2.4739a	2.3717a	2.3521a	2.3415
Gamma/P/sigma			0.0237a	2.3168a			2.3168a	0.683b	0.4277a
Kappa									0.5024a
Gamma variance / lntheta	0.4414a	-17.8290a	-15.094a	-0.5976		-17.8290a	-0.5976	-15.5143a	0.0795
No of observations	18916	18906	18906	18906	18906	18906	18906	18906	18906
Wald test for joint Signif. Of variables	chi2(18)=750.573 Prob>chi2=0.0000	chi2(17)=448.93 Prob>chi2=0.0000	chi2(17)=558.93 Prob>chi2=0.0000	chi2(17)=430.79 Prob>chi2=0.0000	Chi2(17)=570.30 Prob>chi2=0.0000	Chi2(17)=448.93 Prob>chi2=0.0000	Chi2(17)=1006.57 Prob>chi2=0.0000	Chi2(17)=765.11 Prob>chi2=0.0000	
Wald test specification Ho: model well specified	Chi2(1)=119.18 Prob>chi2=0.0000	chi2(1)=1.08 Prob>chi2=0.2987	chi2(1)=1.79 Prob>chi2=0.1810	Chi2(1)=1.04 Prob>chi2=0.3083	Chi2(1)=6.06 Prob>chi2=0.0138	chi2(1)=1.08 Prob>chi2=0.2987	Chi2(1)=1.48 Prob>chi2=0.2235	Chi2(1)=0.41 Prob>chi2=0.5222	
Test proport. Haz. assumption (global test)					Chi2(17)=45.79 Prob>chi2=0.0002				
LR test existence Unobs. Heterogeneity	chi2(1)=76.0975 Prob>chi2=0.0000								
Log-likelihood	-4663.9427	-1719.8305	-1656.7566	-1419.5643	-8715.7555	-1719.8305	-1419.5643	-1478.2255	-1478.6699
AIC	9365.885	3475.661	3351.513	2877.129	17467.51	3475.661	2877.129	2994.451	2997.34

Notes: a, b, and c stand for significance level of 1%, 5% and 10% respectively. Estimations are robust. † - with the Cox function we use the *cluster(patient)* option as a proxy for heterogeneity/frailty since the frailty option required a matrix size higher than the standard one provided by STATA.

Table 5A: Comparison between discrete and continuous time models with interaction terms and accounting for heterogeneity. Coefficients are presented.

	Proportional Hazard Form	**Accelerated Failure-Time Form**
Continuous Payment	**Continuous time Gompertz**	**Continuous time LogNormal**
Hazard	**Coef.**	**Coef.**
Trauma	-1.2234a	0.7406a
Hac	-0.5203a	0.4778a
Rg1	0.1524b	-0.0376
Rg3	-0.1777a	0.2656a
Rg4	-0.4240a	0.4091a
Age	-0.0052	0.0031
Male	-0.1258c	0.0723
Lnincome	0.1426a	-0.1181a
Lnpay1	**-0.0063c**	**0.0200a**
Hacalnpay1	**0.0073a**	**-0.0172a**
Traumaalnpay1	**-0.0712a**	**0.0435a**
Lnpay2	**-0.0969a**	**0.0817a**
Hacalnpay2	**0.0641a**	**-0.0492a**
Traumaalnpay2	**0.0228a**	**-0.0048a**
Student	0.1536	-0.1544c
Unemploy	-0.0391	-0.0366
Privwork	0.1026a	-0.0971a
Selfwork	0.0843	-0.0061
Retired	-0.0518	0.1135
Houswife	-0.0799	0.0779
Exempt	-0.1749	0.0131
_cons	-2.5896a	2.2075a
Gamma/ P/ Sigma	0.0250a	0.6744b
lntheta	15.4525a	-16.7341a
No of observations	18906	18906
Wald test for joint signif. Of variables	chi2(21)=575.19 Prob>chi2=0.0000	Chi2(21)=796.38 Prob>chi2=0.0000
Wald test specification Ho: model well specified	chi2(1)=0.27 Prob>chi2=0.6018	Chi2(1)=3.34 Prob>chi2=0.0676
Log-likelihood	-1642.2553	-1459.5583
AIC	3322.511	2957.117

Notes: a, b, and c stand for significance level of 1%, 5% and 10% respectively. Estimations are robust.

Table 4B Comparison between discrete and continuous time models. Coefficient estimates associated with the payment variables presented by hospital.

	Hospital 1	Hospital 2 - HAC	Hospital 3 - Trauma	
	Payment in the ward	**Payment in the ward**	**Payment in the AD**	**Payment in the ward**
Cloglog PH	-0.1395a - misspec.	-0.0519a - misspec.	-0.0742a - misspec.	-0.0730a - misspec.
Exponential PH	-0.0738a	-0.0278a	-0.0582a	-0.0629a
Gompertz PH	-0.1411a	-0.0542a	-0.0747a	-0.0755a
Weibull PH	-0.1413a	-0.0520a	-0.0806a - misspec.	-0.0825a - misspec.
Cox	-0.1119a	-0.0447a - misspec.	-0.0762a - misspec.	-0.0754a - misspec.
Exponential AFT	0.0738a	0.0278b	0.0582a	0.0629a
Weibull AFT	0.0729a	0.0261b	0.0567a - misspec.	0.0580a - misspec.
LogNormal AFT	0.0748a	0.0312a - misspec.	0.0647a - misspec.	0.0800a - misspec.
Gener. Gamma	0.0746a	0.0293a - misspec.	0.0583a - misspec.	0.0634a - misspec.

Table 5B Comparison between discrete and continuous time models taking heterogeneity into account. Coefficient estimates associated with the payment variables presented by hospital.

	Hospital 1	Hospital 2 - HAC	Hospital 3 - Trauma	
	Payment in the ward	**Payment in the ward**	**Payment in the AD**	**Payment in the ward**
Cloglog PH	-0.2195a - misspec.	-0.1744a - misspec.	-0.0840a - misspec.	-0.0859a - misspec.
Exponential PH	-0.0738a	-0.0278b	-0.0582a	-0.0629a
Gompertz PH	-0.8148	-0.2017a	-0.1842a	-0.2610a
Weibull PH	-0.2205a	-0.0912a - misspec.	-0.1018a	-0.1115a
Cox	-0.1119a	-0.0447a - misspec.	-0.0762a - misspec.	-0.0754a - misspec.
Exponential AFT	0.0738a	0.0278b	0.0582a	0.0629a
Weibull AFT	0.0741a	0.0271a - misspec.	0.0604a	0.0662a
LogNormal AFT	0.0748a	0.0312a - misspec.	0.0647a - misspec.	0.0800a - misspec.
Gener. Gamma	0.0682a - misspec.	0.0287a - misspec.	0.0561 – misspec.	0.0562 – misspec.

Table 6: Comparison between discrete and continuous time models using continuous payment variable (different specifications of the Hazard function). Coefficients (rather than ratios) are presented.

	Proportional Hazard Form					Accelerated Failure-Time Form			
Continuous Payment	Discrete time PH Cloglog	Continuous time Exponential	Continuous time Gompertz	Continuous time Weibull	Continuous time Cox	Continuous time Exponential	Continuous time Weibull	Continuous time LogNormal	Continuous time Gener. gamma
Hazard	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Logd	0.5525a								
Trauma	-1.5552a	-0.9692a	-1.2904a	-1.6299a	-1.3769a	0.9692a	1.0076a	0.8247a	0.9434a
Hac	-0.2933a	-0.2071a	-0.2073a	-0.2950a	-0.3181a	0.2071a	0.1823a	0.2223a	0.2145a
Rg1	0.0971	-0.0054	-0.0079	0.1106	0.0869	0.0054	-0.0683	0.1199b	0.0354
Rg3	-0.4001a	-0.2951a	-0.3241a	-0.3995a	-0.3932a	0.2951a	0.2469a	0.4008a	0.3194a
Rg4	-0.6331a	-0.4531a	-0.5263a	-0.6538a	-0.6460a	0.4531a	0.4042a	0.5372a	0.4743a
Age	-0.0139a	-0.0090a	-0.0108a	-0.0142a	-0.0129a	0.0090a	0.0087a	0.0078a	0.0089a
Male	-0.0997	-0.0378	-0.0449	-0.1026	-0.0904	0.0378	0.0634	-0.0148	0.0261
Lnincome	0.4491a	0.3006a	0.4066a	0.4745a	0.3932a	-0.3006a	-0.2933a	-0.3395a	-0.3073a
Lnpay1	**-0.0843a**	**-0.0585**	**-0.0827b**	**-0.0889c**	**-0.0710c**	**0.0585**	**0.0550**	**0.0723b**	**0.0610c**
Lnpay2_hat	**-0.3296a**	**-0.2255a**	**-0.3146a**	**-0.3478a**	**-0.2842a**	**0.2255a**	**0.2150a**	**0.2718a**	**0.2335a**
Student	-0.0248	-0.0069	-0.0123	-0.0212	-0.0057	0.0069	0.0131	-0.0133	0.0043
Unemploy	-0.1981b	-0.1105	-0.1763	-0.1942	-0.1591	0.1105	0.1201	0.0823	0.1047
Privwork	0.0702	0.0419	0.0373	0.0844	0.0904c	-0.0419	-0.0522	-0.0307c	-0.0370c
Selfwork	0.3507b	0.2026	0.3525c	0.3707	0.2269	-0.2026	-0.2292	-0.2170	-0.1988
Retired	-0.5186a	-0.3556a	-0.46662a	-0.5289a	-0.4727a	0.3556a	0.3269a	0.4676a	0.3759a
Houswife	-0.2385b	-0.1584a	-0.1953a	-0.2386a	-0.2158a	0.1584a	0.1475a	0.1897b	0.1645a
Exempt	-0.0425	-0.01429	-0.0926	-0.0682	0.0127	0.01429	0.0422	-0.0137	0.0045
_cons	-4.3551a	-3.0315a	-3.5985a	-5.0976a		3.0315a	3.1513a	3.0144a	3.0057a
Gamma			0.0244a						
P				1.6176a			1.6176a		
Sigma								0.6893b	0.6471a
Kappa									0.4811a
No of observations	19008	18998	18998	18998	18998	18998	18998	18998	18998
Wald test for joint signif. Of variables	chi2(18)=722.0886 Prob>chi2=0.0000	chi2(17)=430.58 Prob>chi2=0.0000	chi2(17)=542.59 Prob>chi2=0.0000	chi2(17)=825.10 Prob>chi2=0.0000	chi2(17)=693.16 Prob>chi2=0.0000	chi2(17)=430.58 Prob>chi2=0.0000	chi2(17)=1249.84 Prob>chi2=0.0000	chi2(17)=667.96 Prob>chi2=0.0000	chi2(17)=941.13 Prob>chi2=0.0000
Wald test omitted vars Ho: no omitted vars	Chi2(1)=112.99 Prob>chi2=0.0000	chi2(1)=0.04 Prob>chi2=0.8347	chi2(1)=1.79 Prob>chi2=0.1807	chi2(1)=0.20 Prob>chi2=0.6551	chi2(1)=16.28 Prob>chi2=0.0002	chi2(1)=0.04 Prob>chi2=0.8347	Chi2(1)=0.22 Prob>chi2=0.6397	Chi2(1)=0.06 Prob>chi2=0.8016	chi2(1)=0.03 Prob>chi2=0.8542
Test proport. Haz. Assumption (global test)					chi2(17)=16.28 Prob>chi2=0.0001				
Log-likelihood	-4743.4918	-1735.9816	-1671.7592	-1501.5937	-8789.8146	-1735.9816	-1501.5937	-1499.1786	-1463.7692
AIC	9524.984	3507.963	3381.518	3041.187	17615.63	3507.963	3041.187	3036.357	2967.538

Notes: a, b, and c stand for significance level of 1%, 5% and 10% respectively. Estimations are robust. Wald test for kappa=1 (Ho:kappa=1) : Chi2(1)=6.93 and Prob>chi2=0.0085

Figure 1A) crude hazard function scale from 0-1 and the density function for patients' discharge

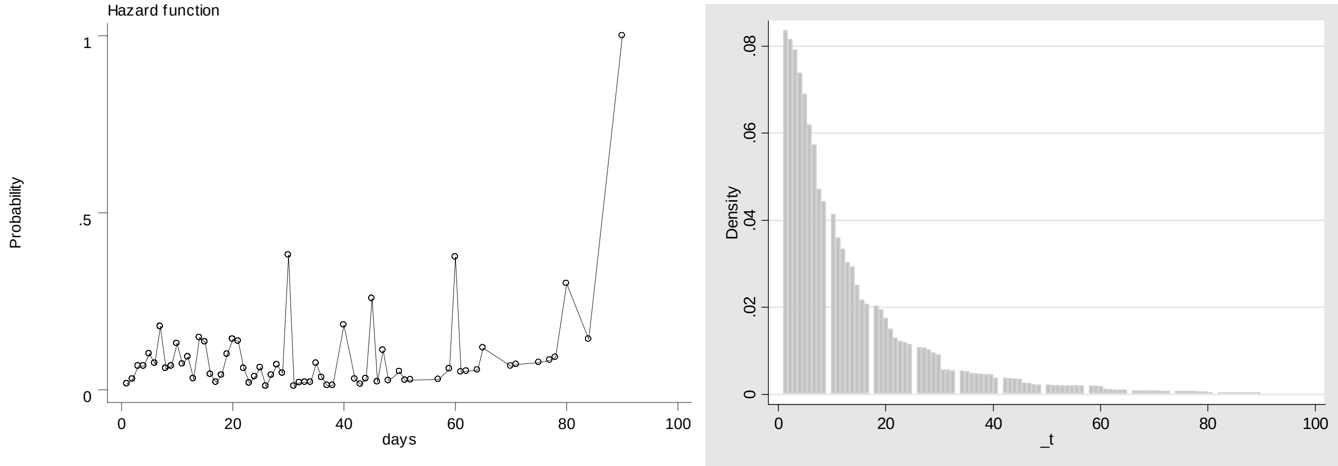

Figure 1B) weighted kernel of the estimated hazard contributions scale from 0-0.08

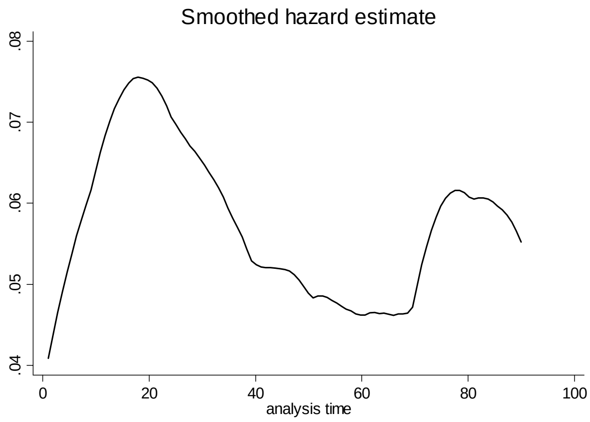

Figure 2: A) Nelson-Aalen cumulative hazard function pooled sample and B by gender

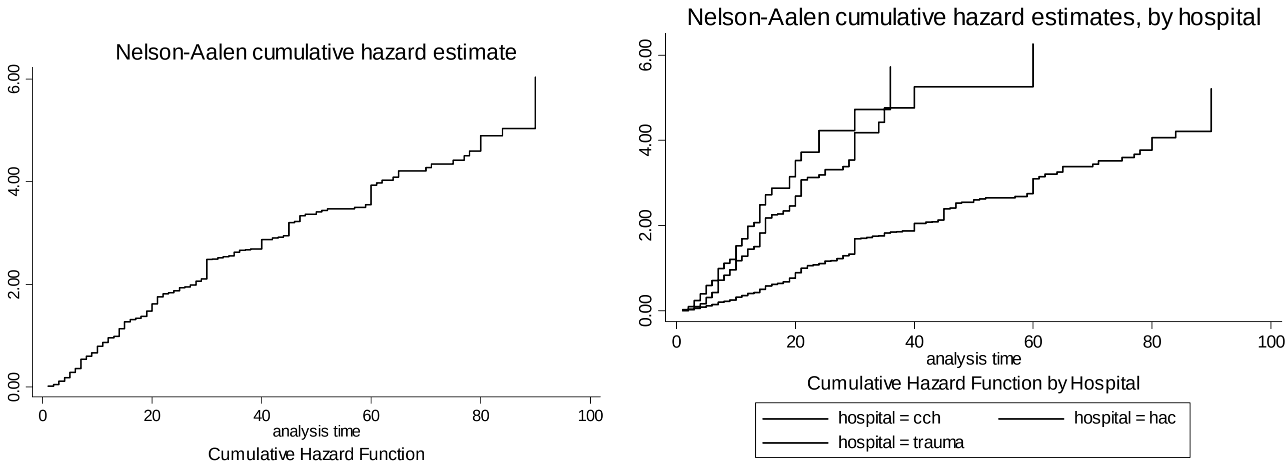

Figure 3: A) Kaplan Meier survival function pooled sample and B by gender