NOVA SCHOOL OF BUSINESS AND ECONOMICS

UNIVERSIDADE NOVA DE LISBOA

Dissertation, presented as part of the requirements for the degree of

Doctor of Philosophy in Economics

A tour of health care:
emergency room, hospital and home

Luís Alexandre Coelho Filipe, Student Number 42

A dissertation carried out on the PhD in Economics, under the supervision of Professor Pedro Pita Barros.

May, 2020

Abstract

My thesis is a small tour of the health care system (emergency room, hospital and home), with special emphasis on providers. Chapter 1 analyzes how emergency room doctors change their behavior when the waiting room is crowded. The outcomes reflect the time spent with the patient, the intensity of treatment, and discharge destination. Chapter 2 extends the previous setting to inpatient care, to determine how doctors react to hospital occupancy level. It identifies doctors' discharging criteria as a causal factor for the positive relation between occupancy rates and readmissions. The analysis in Chapters 1 and 2 contributes to the doctors' incentives literature, explaining how these agents behave in the context of a National Health Service, with no financial incentives. Chapter 3 examines the impact of informally providing care to a partner (at home) on the physical and mental health of the carer.

Keywords: Doctors decision making, queues & length of stay, occupancy rates & readmission episodes, spousal caregivers & caregiver burden.

Acknowledgments

First and foremost I would like to thank my advisor Professor Pedro Pita Barros. Thank you for all the support and encouragement. It was really fun being one of your pupils! I also want to say that I admire your motivation tactics: you make sure your students know you trust them and then they become too afraid to disappoint you. It works indeed...

To Bruno Martins, my amazing co-author and one of my best friends. If there is any problem with the first chapter of this thesis he is the one to blame!

Thanks to Rafael Barbosa, another one of my best friends, and to Bruno Martins (again), for reading the thesis and doing some proofreading. I do not want to sound repetitive, but if there are any typos or English mistakes it is their fault.

Many thanks to Céu Mateus, my career advisor. Céu has been very important during this journey as a PhD student, providing me always the best advise.

Thanks to Susana Peralta, my social advisor. Besides throwing parties at her house in Areia Branca, Susana is always there for PhD students.

Thanks to all the friends I made in ESHP, Erasmus School of Rotterdam, specially to Werner Brouwer and Job van Excel. You were nothing short of amazing in welcoming me to Rotterdam.

I am grateful to all the participants of the Nova Health Economics and Management Knowledge Center and Nova's PhD research group.

I am grateful to *FCT — Fundação para a Ciência e Tecnologia*, for financial support. (reference SFRH/BD 115692/2016)

Thanks to all my friends.

Thanks to all my family, specially my parents and my wife!

I dedicate this thesis to my grandmothers and to my parents. They were always very proud of my achievements as a student. Most of all, I know they are very happy that I made the best out of an opportunity they never had. They worked hard to provide me with that opportunity!

Contents

List of Figures

List of Tables

List of abbreviations

ACSS	Administração Central do Sistema de Saúde, I.P.
ARS	Administrações Regionais de Saúde
C $\rightarrow$ C	Spousal caregivers when first observed who stay spousal caregivers in the next period
DRG	Diagnosis Related Group
ED	Emergency Departments
GLM	Generalized Linear Model
IV	Instrumental Variable
LOS	Length of Stay
LPM	Linear Probability Model
LVT	Lisbon and the Tagus Valley
MDC	Major Diagnostic Category
MTS	*Manchester* Triage System
NC $\rightarrow$ C	Non carers when first observed who become carers in the next period
NHS	National Health Service
OLS	Ordinary Least Squares
OOP	Out-of-Pocket payment
OProbit	Ordered Probit
PPP	Partnership between the state and a private institution
SAH	Self Assessed Health
SHARE	The Survey of Health, Ageing and Retirement in Europe
UK	United Kingdom
ULS	Local Health Units

Introduction

In this thesis, I offer three examples of how providers fare in three distinct stages of the health system. In Chapter 1, I address the response of doctors in emergency care to an increase in patients waiting for treatment. In Chapter 2, I study the relation between occupancy rates and readmission episodes through the optic of doctors' discharging criteria. These two chapters focus on the Portuguese National Health Service where doctor's have no performance financial incentives. In Chapter 3, I examine the possible consequences of spousal informal caregiving on the health outcomes of the providers in Europe.

In Chapter 1, "Doctors' Response to Queues: Evidence from a Portuguese Emergency Department", with Bruno Martins, we analyze how doctors' behavior changes with the queue size of an emergency department. Resources in a fast-care setting such as emergency departments are limited, (namely physicians' time) and so it is important to study the extent to which those restrictions reflect on a lower provision of care when more patients are waiting. We evaluate doctors' response to an increase in the number of people waiting at three different stages. First, whether they spend less time with patients. Second, whether they refer patients for fewer exams and lab tests during their visit. Third, whether doctors are more likely to send patients home or to a primary care facility, rather than treating them in the hospital. We use visit-level data spanning almost two years, from a hospital in Lisbon, to estimate a high-dimensional fixed effects model that exploits variation in the queue size at a very fine year-month-day-hour level. We estimate the causal effect of the queue on doctors' behavior using the number of arrivals in the previous sixty minutes as an instrumental variable for the number of patients waiting. Our results show that, as the queue size increases, doctors spend less time with patients, and prescribe fewer lab tests and exams. At the same time, patients are more likely to be sent to a primary care facility, while less likely to be admitted to the hospital. We also study the heterogeneity of doctors' response to the queue according to patients' urgency level, determined at the emergency department triage. We find heterogeneity in the time spent with the patient, but not in the decision to send patients for exams/lab tests. When the queue increases, physicians shorten the length of stay of patients with low urgency more than those with a very high urgency visit. Rationing both treatment and diagnosis are not bad per se, and whether it has negative health consequences is an empirical question that depends on the health production function. Indeed, in the presence of a concave utility function, marginal returns to health care are decreasing, and, if care is sufficiently high, then rationing will not have serious health effects. It is encouraging that we find some evidence that physicians ration efficiently, in the sense treatment intensity is reduced most for non-urgent patients, whose health impacts are more likely to be small.

I extend the setting in the previous chapter to inpatient care in Chapter 2: "A bed constraint? Hospital occupancy rates and readmission episodes". In this chapter, I focus on how admitted patients are affected by resources constraints. I look at the relationship between occupancy rates and readmissions to implicitly assess doctors' decision-making process when hospitals are more crowded. I explore the hypothesis that doctor's react to time and resources constraints by early discharging patients, due to pressure to admit incoming sicker patients. I use patient level data and a wide set of fixed effects, at the year-month-week-hospital level and DRG-severity level (Diagnosis Related Group), to conclude that occupancy rates and variation in occupancy rates just prior to discharge are predictors of future readmission episodes. Then, I use exogenous length of stay intervals defined in the DRG national tariffs list to interact with occupancy rates. Results show that when occupancy rates are higher, patients discharged earlier than expected are more likely to be readmitted in the future. This result establishes the relation between early discharges and readmissions. I also find heterogeneity across age groups. The most affected by occupancy rates are the elderly, particularly the elderly who are discharged earlier than expected. This fact constitutes a relevant source of inequality.

Chapter 3, “Partners in care! Health effects of providing care to spouses or partners”, examines the impact of providing care to a partner on the physical and mental health of the carer. Using individual level data from five pooled waves of the SHARE survey (The Survey of Health, Ageing and Retirement in Europe), I start by running models of univariate regression analysis on outcomes of physical health and mental health, and observe that if selection is not accounted for, then carers will always show lower physical and mental health in comparison to partners without caring responsibilities. To address this issue, I include patient characteristics in the model, to account for emotional burden, correlated behaviors and household health investment decisions, and there is a dramatic change in the results. When models account for the patient characteristics, the physical health of partners with caring responsibilities is better than the one of partners without a caring role. However, I do no observe any meaningful differences for mental health.

Then, I perform propensity score matching to ensure that both groups have potential patients with similar levels of health. The results lead to the same conclusion; carers show a slightly better physical health status. This methodology does not guarantee causality, as for instance, non-carers' physical condition may deter them from aiding their partners.

To account for that, I perform an event study that compares individuals who transition into caregiving with the ones who remain non-caregivers across waves of the survey. I match observations of both groups in the pre-treatment period, on individual and partner characteristics, to improve the likelihood that the transition is random. The results of becoming a caregiver are positive for self perceived physical health, and non-conclusive for mental health and disability.

Part I

Emergency care

Chapter 1

Doctors' Response to Queues: Evidence from a Portuguese Emergency Department[1]

1.1 Introduction

Hospital Emergency Departments (ED) are known for consistently operating at over capacity, with long queues leading to waiting times that can last several hours. Patients' arrivals to the ED pose an externality onto other patients since they decrease the amount of resources available. This externality can potentially lead to negative health outcomes for patients, either because their medical condition worsens while waiting, or because there are fewer resources available for treatment/diagnosis. Policy-makers and managers have the difficult task of assigning capacity limits (in terms of physical and human capital) to hospitals, requiring them to balance the potential health consequences of peak-load times and the opportunity cost of capital under idle times.

In this paper we investigate how physicians respond to the level of visits in the ED, in the context

[1]Co-authored by Bruno Martins, PhD in Economics from Boston University

of a Portuguese public hospital. We study responses in three different aspects. First, we analyze the time spent during visits, i.e., length of stay (defined as the time between being called to see the doctor for the first time and being discharged from treatment). Second, we look at the intensity of the use of lab tests, exams, and treatment procedures (which we proxy with the out-of-pocket payment (OOP) due to the institutional context of our setting). Third, we study the doctors' choice of destination upon discharging patients. We cannot directly study patient outcomes, but we study the effect of the queue on waiting times - time between finishing the triage and being called to see a doctor - which have been shown to be correlated with negative patient outcomes (Hoot and Aronsky 2008).

We use visit-level data for one ED in the Lisbon area, between January 2011 and October 2012, to construct the waiting queue for an individual patient based on timestamps for several milestones within an ED visit. We leverage on a fixed-effects model to exploit very fine variations in queue size occurring within a specific hour, thus controlling flexibly for within-day variation, as well as weekly and monthly seasonality and yearly trends. We control for the endogeneity of the queue by using the number of arrivals to the ED in the past 60 minutes as an instrument. Visits to the ED have been used in the literature as a source of exogenous variation in health expenditures (Eichner 1998, Kowalski 2016).

Literature on doctors' behavior mostly studies the supply of medical care under different payment mechanisms, based on the seminal work by Ellis and McGuire (1986).[2] With no financial incentives, physicians' degrees of agency depend roughly on sense of duty, altruism, (Kolstad 2013, Godager and Wiesen 2013) and peer or patient pressure (Chan 2016, Silver 2016).

Time constraints drive doctors to optimize based on a time-quality trade-off (Anand, Paç and Veeraraghavan 2011, Dugdale, Epstein and Pantilat 1999). These choices are important, as time spent with patients is related to positive health outcomes (Ogden et al. 2004, Chen, Farwell and Jha 2009,

[2] A brief overview of policy reforms that change physicians payment mechanisms can be found in Chandra, Cutler and Song (2011) or McClellan (2011).

Silver 2016) while time waiting for treatment is associated with undesirable events (Sivey 2018, Baker, Stevens and Brook 1991, Bindman et al. 1991).

We contribute to the literature by empirically testing the hypothesis that doctors are aware of and react to the number of patients waiting in the ED. Patients in an ED wait for care at the facility, thereby creating extra pressure on doctors to react to congestion. Also, due to the fixed physician compensation, Portuguese ED provide a perfect setting to study these decisions, net of financial incentives. The literature so far has focused on waiting lists where this pressure does not exist, such as elective surgery (Riganti, Siciliani and Fiorio 2017) or elective admissions (Siciliani, Stanciole and Jacobs 2009). Our measure of waiting time differs from these in that our time horizon is much shorter, and measured in hours instead of days.

Our work is closely related to Sivey (2018) and Gruber, Hoe and Stoye (2018). Sivey (2018) evaluates the consequences of waiting times in Australian emergency departments on the demand-side, studying the likelihood of patients exiting the waiting room without treatment when waiting times increase. Our paper is a complement, in the sense that we use a similar setting but rather answer the supply-side question of doctors' behavior. Gruber, Hoe and Stoye (2018) study how waiting times in emergency departments in the UK lead to changes in the probability of admission and mortality. They show that a decrease in waiting times leads to a decrease in mortality. Gruber, Hoe and Stoye (2018) establish that doctor's actions have serious consequences for patients' health status (a number of other studies have correlated waiting times with patient mortality, as summarized by Hoot and Aronsky 2008). Therefore, while we do not have data to assess mortality, evaluating doctors' behavior has important consequences. In our analysis, we show the relationship between queues and waiting times, to assess whether there is a channel through which physicians attempt to decrease waiting times, and consequently the associated negative outcomes.

Our results show that doctors react to increased visits to the ED in several dimensions. First, they spend less time with patients. A 1% increase in the number of patients waiting leads to a 0.42% decrease in length of stay. Furthermore, there is evidence of externalities across patients of different urgency degrees, since doctors decrease patients' length of stay when the queue of a different urgency level increases, suggesting that patients put pressure on resources across the board. Second, doctors decrease costs of diagnosis and treatment (implying lower intensity), driven by the extensive margin. Finally, doctors are less likely to admit patients to inpatient care.

We find mixed evidence of heterogeneity in the impact by arrival urgency. An increase in the number of visits leads to a greater decrease in the length of stay in the ED for patients who had a non-urgent visit than for those with a very urgent visit. However, this heterogeneity is not present in terms of costs of treatment and diagnosis.

Understanding doctors' reactions to queues in emergency departments is the first step for policy makers to design mechanisms aimed at achieving optimal outcomes. For example, policy makers seeking to reduce congestion face important indirect costs. Our results suggest that doctors use more resources per patient as queues decrease, which could increase the financial costs for payers, such as in a National Health Service. In addition, our findings also suggest a positive correlation between the utilization of hospitals and primary care centers. A peak in admissions to hospitals increases admissions at primary care centers, as doctors substitute their patients between the two. Therefore, choosing capacity for each type of care facility must be a joint decision that accounts for the positive correlation.

In Section 3.2 we describe the data and how we define our main variables of interest. In Section 1.3 we present our empirical strategy and discuss identification. Section 3.4 presents our main results, while Section 1.5 focuses on robustness checks to further support our findings. Finally, we conclude in Section 3.6.

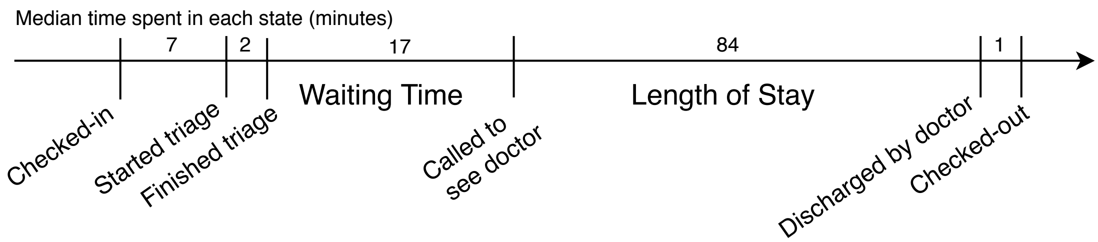
1.2 Data and Summary Statistics

We make use of visits-level data of a single hospital in the Lisbon area between January 2011 and October 2012. Appendix A.1 provides a brief description of the Portuguese health system.

All visits occurring within that period are recorded electronically in the hospital information system.

For each visit we observe several time stamps: the times (detailed to the second) in which a patient checks in, starts, and finishes the triage, is called to be seen by a doctor for the first time, is discharged by the doctor, and checks out from the ED. We present a schematic timeline in Figure 1.1, along with the median time patients spend in each phase.

Figure 1.1: Arrivals time-line



Note: Figure shows a typical visit to the emergency department, divided into the several milestones that are present in the data.

As seen in Figure 1.1, we define waiting time as the time spent in the hospital after the triage and before being called to see the doctor [3]. We define patient length of stay (LOS) as the time between the patient being summoned to see a doctor for the first time until she gets discharged.[4] Note that our measure of length of stay does not include the waiting time, and focuses only on the time spent after waiting to be called for treatment.

Upon being called from the waiting room, the patient may see a nurse (to take the temperature, for example). After being seen by a doctor, the patient may be sent for further testing (such as an X-Ray). Then, the patient needs to see the doctor again before being discharged. Our measure of length of stay is a combination of all those factors. Length of stay is a term usually applied to time spent in inpatient care. However, in our framework it represents the time patients spend under treatment/diagnosis in an outpatient setting. This notation is the same used in Chan (2016).

[3] We chose not to include the time waiting before the triage because at that point in time, the urgency of the visit would not have been defined.

[4] We do not observe any other time stamps within that interval.

Our dataset has results of the triage, using the *Manchester* Triage System (MTS), in which each patient is given a color code that represents the *ex-ante* urgency of the visit. There are five color codes, which can be seen in Table 1.1, ranked into levels of increasing urgency.[5] The MTS is an algorithm that triage nurses follow in order to determine the urgency of the visit. The algorithm includes major symptoms and further narrows down additional signs ranked by priority. Zachariasse et al. (2017) show, using information from a Lisbon-area hospital, that the MTS in Portugal is a good, but not perfect, predictor of urgency level.

In Table 1.1, as well as in Figure A.1, we describe our main dependent variables. The table shows that waiting times are decreasing in urgency, while out of pocket is increasing. Length of stay is also increasing from green to orange. But it decreases in red urgency level visits as those patients might be sent to the intensive care unit or die in the ED. The median value for the length of stay is below the mean, revealing the data to be skewed to the right. To account for this, we analyze this variable in logarithmic terms.

Length of stay in our data includes a waiting period that patients face in between treatment/diagnosis stages. This period is weakly positively correlated with visits to the ED, as more people are waiting for additional testing. Therefore, a higher number of arrivals to the ED might increase total length of stay, even if doctors decrease time spent with the patient. For this reason, we expect our estimate to be a lower bound on the true doctors' response.

To evaluate doctors' decisions regarding the intensity of treatment, tests, and exams when facing longer queues, we use patient out-of-pocket payments.[6] This variable is used given the absence of data about the services the patients benefited from. The OOP is increasing with the costs of procedures, which makes it a good proxy for doctors' intensity of treatment and diagnosis instruments.

[5]There is also an additional color code, white, that corresponds to patients who were referred from another doctor and, therefore, the color does not directly represent urgency. We exclude these referrals, which represent 3.2% of our sample.

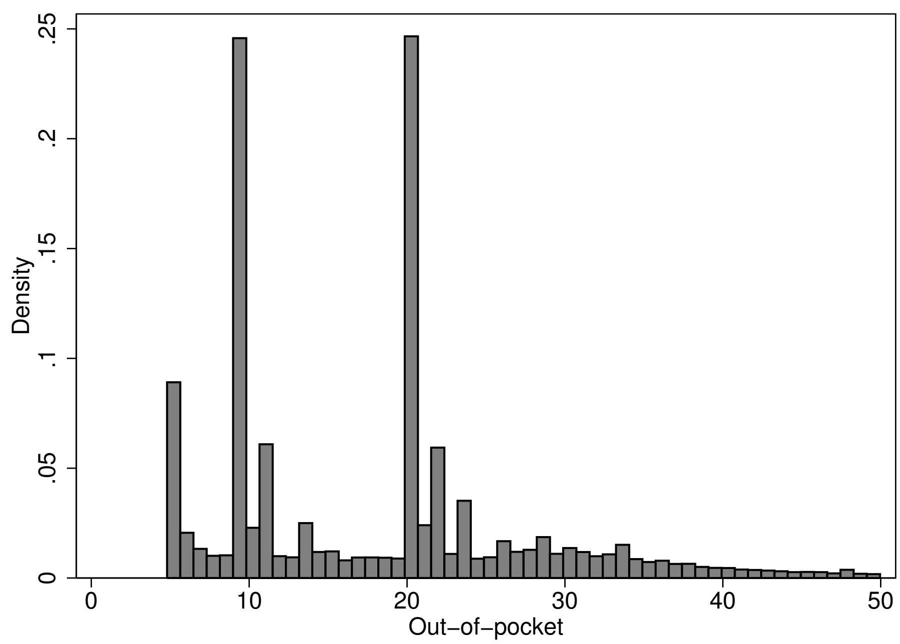
[6]We observe the out-of-pocket the patient was charged regardless of whether she paid or not.

Table 1.1: Summary Statistics

		Blue	Green	Yellow	Orange	Red	All
Observations		3,098	151,980	103,117	16,710	1,374	276,279
Length of Stay	Mean	139.06	128.90	274.56	465.82	353.49	204.87
	SD	951.31	345.38	539.33	737.63	644.57	478.08
	Median	32.68	50.27	139.30	270.07	126.72	84.62
Waiting time	Mean	103.58	62.59	42.67	21.64	20.83	52.93
	SD	140.67	135.96	88.62	24.30	26.98	116.38
	Median	46.89	35.37	28.07	16.47	15.38	30.37
Out of Pocket	Mean	14.23	16.50	21.88	28.97	37.72	19.34
	SD	8.27	9.68	14.67	21.38	42.17	13.55
	Median	9.60	15.55	20.00	23.80	21.21	20.00
Queue	Mean	12.13	14.38	14.30	14.15	15.00	14.58
	SD	9.96	10.84	11.14	10.67	11.17	10.98
	Median	9.00	12.00	12.00	12.00	13.00	12.00

Note: Triage colors are sorted by increasing urgency level ranking from left to right. Length of stay is the number of minutes between being first seen by the doctor and discharge; OOP is the out-of-pocket payment of the visit to the ED, in Euros; Queue is the number of people in the waiting room at the time patients are called to see the doctor. This table uses a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012.

Figure 1.2: Out-of-pocket distribution

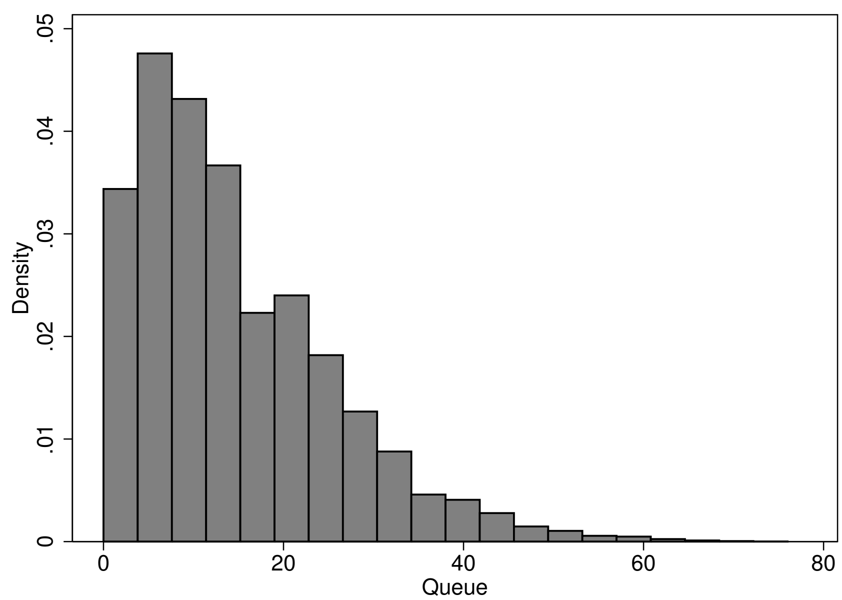


Note: The out-of-pocket is the amount, in euros, that a patient is charged upon discharge from the hospital ED. The bunching points represent the value of the access charge (the minimum valued paid to access the ED services). The values are 9.6 € in 2011, and 20 € in 2012. Values below 9.6 € correspond to a 50% discount enjoyed by patients over 65 years old in 2011. Values are truncated at 50 € for readability.

We show in Figure 1.2 the distribution of the OOP variable. The three bunching points correspond to the value of the access charge. In 2011, this was 9.6 €, although patients over 65 years old receive a 50% discount, which corresponds to the 4.8 € bunching point seen in the figure. In 2012, the access charge was 20 €. Deviations from this amount correspond to extra lab tests, exams, or treatments that patients underwent during their visit to the ED.

We use the time stamps of each arrival to define the *queue*. For each visit we create a tally of the total number of patients waiting in the ED at the time the patient is called to see the doctor for the first time. The assumption on doctors' behavior is that this information is available to them, that is, doctors are able to estimate how many people are waiting at a given point in time.

Figure 1.3: Queue distribution

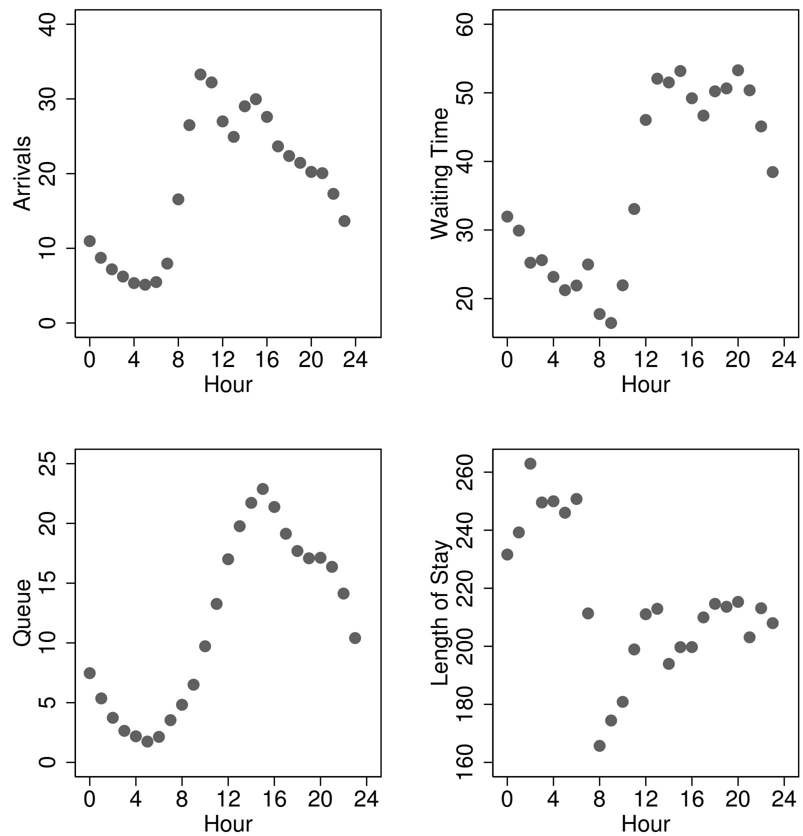


Note: Queue is defined as the number of people waiting in the emergency department at the time patients are called for treatment/diagnosis for the first time.

The distribution of the overall congestion variable can be seen in Figure 1.3. Figure A.2, in the online appendix, shows the distribution of the color-specific queue separately for each type of visit as measured by the triage. More specifically, each panel uses a subset of the visits based on their triage color and shows the queue distribution of the *same* color. For example, the panel labeled

"Blue" takes all the visits with a Blue-level urgency and shows the distribution of the number of people waiting that also had a Blue urgency. The figure shows that there is hardly any congestion for cases that are very urgent (orange and red).

Figure 1.4: Hourly distribution of main variables



Note: Arrivals correspond to the number of patients arriving at the hospital. Waiting time is the number of minutes between finishing triage and being called to see the doctor. Length of stay is the number of minutes between being first seen by the doctor and discharge. Queue is defined as the number of people waiting in the emergency department at the time patients are called for treatment/diagnosis for the first time. Y-axis shows the hourly mean of each variable.

Figure 1.4 provides further statistics, showing the daily distributions of arrivals, waiting times, queues, and length of stay. We observe clear distinctions between night and day, as both demand and supply factors change significantly across both periods. Table A.1 shows the summary statistics separated by color urgency level.

Finally, we observe age and gender of the patients, which allows to control for patient characteristics known to be correlated with general health status. In our sample the average age is 53.66 years old and the percentage of female patients is 53.21%.

Unfortunately, we do not observe any supply variables. Yet, our identification relies on using year-month-day-hour fixed effects to exploit only variation in a single hour, in which the number of doctors is fixed, and which controls for selection of doctors into shifts at different hours.

1.3 Methodology

1.3.1 Length of Stay

Our analysis seeks to evaluate the impact that the ED queue has on the doctors' choice of length of stay. To do this, we start with the following basic specification.

$$\text{Length of Stay}_{ist} = \beta \text{Queue}_{it} + \gamma X_{ist} + \tau_t + \varepsilon_{ist} \tag{1.1}$$

where Length of Stay$_{ist}$ is the total time that patient in visit i, with urgency degree s, in time (hour) t spends on treatment, and Queue$_{it}$ is a measure of how crowded the ED waiting room is. The vector X_{it} includes a set of patient characteristics at the time of visit i, comprising age, gender, age-gender interaction, and urgency level of the visit as measured by the triage color. We include the interaction of fixed effects at the level of year-month-day-hour, τ_t. Therefore, we are exploiting a very fine level of variation in congestion occurring at a specific hour across almost two years of data. Finally, ε_{it} is the idiosyncratic error.

We are primarily interested in the coefficient β, which measures the impact of an increase in the queue on the time spent under treatment. We defined visit i's queue as the number of patients waiting to be seen at the time the patient is called to see the doctor for the first time. Since patients' waiting time often exceeds one hour, the queue for a given patient includes people who arrived at a different hour in the past. Moreover, people who are called to see the doctor in the same hour may have different queues due to, for example, arrivals between the minute in which they were called.

Equation (1.1) implicitly assumes that an additional patient waiting affects all other patients in the same way. However, it is natural to conceive that patients with mild complaints do not cause strong externalities to patients who arrive with life-threatening conditions, since they are likely to use different resources. Therefore, we estimate a second class of models with two main independent variables of interest. First, the queue of patients waiting who have the same level of urgency, as measured by the *Manchester* Triage System; second, the queue of patients who do not have the same triage color.

1.3.2 Identification

Causal interpretation of the effect of arrivals on the outcome variable requires conditional orthogonality between the error term and queue variable. In our case identification may be jeopardized by both demand-side (patients) and supply-side (doctors) factors.

On the demand side, the fact that patients often bypass gatekeepers to obtain specialty care in favor of ED visits implies that these visits are not exogenous - at least for less urgent ones. In particular, the time of day is highly correlated with the number of visits, in the sense that patients are more likely to go to the emergency department in the morning and less likely at night. To prevent this from confounding our estimates, we rely on the very finely defined fixed effects at the year-month-day-hour, which control for selection throughout the day and exploit only variation within a single hour.[7] On the supply side, the cyclicality of treatment throughout the day reflects on the number of doctors available and hours worked in the ED in a given hour. Swami et al. (2018) show that the number of hours worked by general practitioners in primary care in Australia has effects on

[7]This approach does not control for selection within an hour, which we assume to be minimal.

waiting times. While we do not have information regarding the number of doctors in the ED at each time, work shifts do not change within an hour, and so this is also controlled for using the hour fixed effects.

Equation (1.1) also suffers from a reverse causality issue caused by doctors' actions. Less time spent undergoing treatment/diagnosis leads to faster discharges from the ED, which in turn cause a decrease in the number of patients in the waiting room. For this reason, estimating Equation (1.1) by OLS will underestimate the impact of queue size on length of stay.

In order to overcome this issue we use an instrumental variable strategy. We instrument the queue variable with the number of patients who visited the ED in the 60 minutes prior to a patient's arrival. For this instrument to be valid we need arrival of patients to the ED to affect length of stay only by increasing the queue size. This argument requires patients' decisions to go to the ED to be independent from the total visit time. There is no formal mechanism for patients to have real time information about visit times and so they are not likely to be aware of and react to them. [8]

We choose an interval of 60 minutes for the instrument to guarantee a sufficiently high correlation between the instrument and the dependent variables for the most power-demanding specification in which we separately analyze the results by triage color, in which observations for some colors are scarce. In Section 1.5 we conduct a series of robustness checks for instruments that use different time intervals as well as over-identified models.

1.3.3 Heterogeneity by Episode Degree of Urgency

A doctor that decides time to spend with two patients optimally equates the marginal benefit of the two. If these two patients have different benefit functions, then the doctor allocates resources unevenly. This leads to a differential change in treatment intensity given an exogenous need for re-optimizing (an increase in the number of arrivals might change doctors' opportunity cost of

[8] Since 2016, patients can visit the website http://tempos.min-saude.pt to obtain real time information about waiting times for hospitals ED, but this platform did not exist for the years corresponding to our data span.

spending time with a given patient).

We postulate that the color code of the triage system picks up different benefit functions of treatment for each patient. Therefore, we allow doctors to react differently to episodes that fall into different levels of triage, s, by running the following specification.

$$\text{Length of Stay}_{it} = \beta_s \text{Queue}_{it} + \gamma X_{ist} + \tau_t + \varepsilon_{ist} \tag{1.2}$$

This model estimates a more demanding specification in which we estimate a β_s for each triage color, while Equation (1.1) averages out these effects into a single coefficient.

1.3.4 Diagnosis and Treatment

In order to analyze doctors' responses regarding diagnosis and treatment procedures, such as lab tests, exams, and drugs administration, we explore the payment scheme of the Portuguese ED system at two different levels. First, the extensive margin, which evaluates the likelihood of a patient being sent to further care after seeing a doctor. Second, the intensive margin, which evaluates the extent to which doctors change the quantity of treatment, conditional on the patient being sent for additional testing.

In a visit to the ED, a patient that paid more than the access charge had additional testing/treatment done. We explore this feature to estimate the extensive margin. We use a linear probability model to estimate the probability that a given patient is sent for further care (example: paid more than 20 € in 2012) after seeing the doctor, using the same right-hand-side variables as in Equation (1.1).

In order to analyze the intensive margin, we use the fact that in 2012 the OOP increased for every extra exam, lab test, or treatment that the patient underwent on her visit. Costlier procedures are mapped into weakly higher OOP payments.[9] In 2011 this is not necessarily the case.[10] There-

[9]For example, procedures that cost between 60 € and 64.99 € have an out-of-pocket of 12 € while procedures that cost between 65 € and 69.99 € have an out-of-pocket of 13 €). The OOP fee schedule is available in Portaria n. 306-A/2011.

[10]The details are available in Portaria n. 34/2009.

fore, we use only 2012 data to analyze the intensive margin. We estimate a conditional model by selecting only patients in 2012 that had an OOP over 20 €. Using these patients, we estimate Equation (1.1) where the left-hand-side variable is the value of the OOP, which measures intensity (the combination of quantity and cost) of the treatment above the access charge.

1.3.5 Destination Target

The last channel of doctors' behavior we explore is the destination of the patient upon discharge. The last decision that doctors face is where to send patients after seeing them. This channel may also be influenced by the number of patients arriving at the ED, in the sense that doctors may cope with a peak in arrivals by sending them through other channels of health care provision. We evaluate this in two steps. First, we assess the probability of admitting the patient to inpatient care versus discharging the patient. Second, for patients who were discharged, we compare the probability of being sent home against the probability of being discharged to be followed up in the primary care services. These three cases account for roughly 80% of the total sample.[11] In each case, the independent variable is a binary indicator for which we run a linear probability model of the same specification as Equation (1.1).

1.3.6 Waiting times

Doctors' choices of treatment intensity when faced with increases in the queue size have an effect on the time new patients spend waiting. If doctors decrease length of stay of patients because of the increasing queue size, then waiting times respond less than one-to-one to the queue size. This is especially important because waiting times have been shown to be correlated with negative outcomes (Sivey 2018) and acting to decrease those times mitigates the associated problems.

[11] 9% are admitted to inpatient care, 31% are sent to primary care, and 39% are discharged home. We leave out options such as being sent to another hospital (6%), outpatient visit in the same hospital (9%), leaving without being seen by the doctor or against doctor's recommendation (6%).

We analyze the impact of the queue on waiting times by running regressions of the same configuration as Equation (1.1), but in which the dependent variable is the time spent waiting for care. Specifically, waiting time is measured as the time between finishing the triage and being called to see a doctor for the first time. Moreover, the queue used in this particular regression is measured as the number of people waiting to see a doctor at the time of check in (as opposed to previous regressions). Therefore, we measure how the number of people waiting in the emergency room at the time of check-in affects the amount of time the patient needs to wait to be called for care.

1.4 Results

1.4.1 First stage

We start by showing the strength of our instrument in Table 1.2. In the subsequent regression analysis, the number of observations is less than in Table 1.1 because we remove from the sample visits in the first 12 hours, as we cannot accurately calculate the queue for the initial observations in our data. The number of arrivals in the previous 60 minutes is a good predictor of the queue, defined as the number of patients waiting to be seen at the time of starting treatment. In fact, the model predicts that an additional patient arriving at the ED in the past hour increases the number of patients waiting by 0.136 when including the fixed effects. When focusing only on episodes that have the same urgency level, the coefficient increases to 0.26, while a unit increase in different urgency increases the corresponding queue by 0.22 (columns 2 and 3). Note also that the cross impact between same and different colors are negative, albeit with a much smaller coefficient. In order to use arrivals in the previous 60 minutes as an instrument, the sign of the coefficients does not matter as long as there is predictive power.

In all three cases the excluded instruments are statistically significant. The F-stats for each of the first-stage regressions are 30.89 for overall visits, 33.85 for visits of the same urgency level, and

52.44 for visits of different urgency level.

Table 1.2: First stage regressions

Dependent variable:	Overall queue	Same urgency queue	Diff. urgency queue
Number of admis. in previous 60 min.	0.136***		
	(0.024)		
Number of admis. in previous 60 min. - Same color		0.263***	-0.098**
		(0.021)	(0.044)
Number of admis. in previous 60 min. - Diff. color		-0.107***	0.217***
		(0.038)	(0.027)
Blue	-0.161**	0.024	-0.140*
	(0.068)	(0.028)	(0.083)
Green	-0.056**	0.448***	-0.529***
	(0.025)	(0.048)	(0.052)
Orange	-0.024	-0.139***	0.150***
	(0.026)	(0.029)	(0.038)
Red	0.001	-0.107***	0.148***
	(0.026)	(0.041)	(0.049)
Female	-0.018**	-0.012*	-0.007
	(0.008)	(0.006)	(0.005)
Age	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)
Female*Age	0.000**	0.000*	0.000
	(0.000)	(0.000)	(0.000)
Instrument F-Stat	30.89	33.85	52.44

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The table presents the effect of the number of persons arriving in the 60 minutes prior to the patient being called by a doctor in the size of the waiting room queue. The three columns make a distinction between overall, same urgency and different urgency queues, respectively. Results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level (15,801 levels). The benchmark triage color is yellow.

1.4.2 Doctors' Response to the Queue

The results of estimating Equation (1.1) are reported in Table 1.3, which shows the impact of the queue on the time spent under treatment.

Table 1.3: Main estimates on the Length of Stay

Dependent variable: Log (LOS)				
	OLS	IV	OLS	IV
Queue/10	-0.091**	-0.286***		
	(0.036)	(0.061)		
Same urgency queue/10			-0.106**	-0.386***
			(0.044)	(0.058)
Different urgency queue/10			-0.077**	-0.183**
			(0.030)	(0.078)
Blue	-1.266***	-1.297***	-1.271***	-1.337***
	(0.036)	(0.032)	(0.035)	(0.034)
Green	-0.876***	-0.886***	-0.857***	-0.756***
	(0.016)	(0.013)	(0.024)	(0.031)
Orange	0.657***	0.652***	0.646***	0.577***
	(0.011)	(0.012)	(0.014)	(0.023)
Red	-0.368***	-0.368***	-0.379***	-0.445***
	(0.092)	(0.084)	(0.096)	(0.089)
Female	0.070***	0.066***	0.070***	0.066***
	(0.018)	(0.017)	(0.018)	(0.017)
Age	0.015***	0.015***	0.015***	0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The table shows the results for the effects of queues on length of stay, making a distinction between overall, same urgency, and different urgency queues. Results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level (15,801 levels). The benchmark triage color is yellow. First stage SW F-stats in Table 1.2.

The estimates for the overall queue suggest that an additional patient waiting to be seen in the ED when the patient is called to see the doctor decreases length of stay by 2.9%. This is a large effect, as for an average length of stay of 205 minutes, it takes only 5 additional patients to cause a decrease in length of stay of almost half an hour. At the average number of patients waiting (14.6), this corresponds to an elasticity of length of stay with respect to the queue of -0.42. Note, however, that despite our linear approximation, the queue might have a non-linear impact on the length of stay (e.g. the effect can be stronger if the queue is larger). When looking only at the number of patients waiting with the same urgency level, a unit increase decreases length of stay by 3.9%. The coefficient on the tally of patients that do not have the same degree of urgency is also negative. However, its magnitude is small compared to the same urgency level. The model rejects the hypothesis that these two coefficients are equal. This result suggests that patients with different magnitudes of arrival urgency do not use the same type of resources and, therefore, impose a smaller externality onto each other.

In Table A.2 in the appendix, we show our results for different choices of fixed effects using the instrumental variables estimation. In all different specifications, the coefficients are negative, but much closer to zero than our benchmark. The estimates show the importance of using the finely detailed fixed effects, as they control for important omitted supply- and demand- side factors, as explained in Section 1.3.2.

We now test whether doctors respond by changing intensity of diagnosis and treatment, at both the extensive and intensive margin.

Table 1.4 shows the results of the instrumental variables estimates on the OOP. The extensive margin, in the first column, shows that an increase in the queue by 1 individual decreases the likelihood of being sent for further care by 0.79 percentage points. In our sample, the likelihood of being sent for further care or diagnosis is 54%, and so an increase in the queue decreases the likelihood by 1.5%, at mean values. Again, patients who are waiting with the same urgency level have a stronger impact on the likelihood of further testing than patients waiting with a different urgency level. As far as the intensive margin is concerned, where we estimate the impact of the queue on the total OOP conditional on the OOP being above the access charge, the coefficient on the overall queue is negative, but imprecise. When disentangling between the queue of the same and different urgency levels, we still find that the impact is driven mainly by the queue of patients who have the same urgency, rather than those of different urgency. Overall, we find evidence that doctors decrease the intensity of treatment when the queue increases, and this is driven mainly by the extensive margin.

Female patients have both lower LOS and OOP when compared to men. Young males might be more likely to end up in the ED injured through participation in sports, while young females might be more likely to resort to emergency care because of anorexia related issues, for example.

1.4.3 Heterogeneous Response

We now turn to the estimation of Equation (1.2), in which we explore the hypothesis that doctors respond to congestion differently for patients with different urgency levels.

We show in Table 1.5 the estimated coefficient for each triage color, measured as deviations from visits with an urgency level of Yellow. An intuitive illustration of the results is displayed in Figure A.3, which shows the differential impact of the queue on each triage color.

The point estimates suggest heterogeneity in doctors' response to overall queues,[12] where we see that less urgent episodes have their length of stay decrease by more than those that are more urgent. For example, an increase in the queue by one patient decreases the length of stay for yellow-colored episodes by 2.6%, while blue-colored episodes have a decrease in length of stay of 4.6%. Besides being small in magnitude, the estimated coefficient for "red" cannot be statistically distinguished from zero, suggesting that increases in congestion do not affect visits that are very life-threatening. Note, however, that only the "blue", "green"[13] and "red" coefficients are distinguishable from each

[12] This section shows the heterogeneous response to the total queue. Results for within-color effects are in Table A.3.

[13] "Green" is statistically distinguishable from "blue" and "red"only at a 90% confidence level. At 95% confidence

Table 1.4: Main estimates on treatment intensity

	Extensive Margin		Intensive Margin	
Dependent variable:	1 if OOP > access charge		Log(OOP|OOP > access charge)	
Queue/10	-0.079***		-0.036	
	(0.017)		(0.024)	
Same urgency queue/10		-0.100***		-0.055**
		(0.016)		(0.025)
Different urgency queue/10		-0.057***		-0.021
		(0.020)		(0.025)
Blue	-0.478***	-0.486***	-0.191***	-0.198***
	(0.010)	(0.011)	(0.013)	(0.013)
Green	-0.239***	-0.212***	-0.161***	-0.142***
	(0.004)	(0.009)	(0.004)	(0.006)
Orange	0.178***	0.163***	0.168***	0.154***
	(0.004)	(0.006)	(0.005)	(0.005)
Red	-0.046***	-0.062***	0.466***	0.452***
	(0.015)	(0.017)	(0.025)	(0.026)
Female	-0.009	-0.009	-0.028***	-0.028***
	(0.006)	(0.006)	(0.006)	(0.006)
Age	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*Age	0.001***	0.001***	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	276,061	276,061	73,609	73,609

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results for the effects of queues on the intensity of treatment, making a distinction between overall, same urgency, and different urgency queues. The extensive margin uses 276,061 arrivals from Jan 2011-Oct 2012. The intensive margin uses 73,609 arrivals from Jan 2012-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level and the queue is instrumented with the number of arrivals in the past 60 minutes. The benchmark triage color is yellow. First stage SW F-stats: 30.89 (column 1); 33.85 and 52.44 (column 2); 32.54 (column 3); 34.35 and 44.59 (column 4).

Table 1.5: Color Specific regressions

Dependent variable:	Log(Length of Stay)		Extensive Margin 1 if OOP > access charge		Intensive Margin Log(OOP|OOP >) access charge)	
	OLS	IV	OLS	IV	OLS	IV
Queue/10	-0.084**	-0.256***	-0.010	-0.069***	-0.009**	-0.038
	(0.031)	(0.066)	(0.006)	(0.017)	(0.004)	(0.025)
Queue/10 × Blue	-0.046	-0.203**	-0.006	-0.034**	0.006	0.041**
	(0.036)	(0.081)	(0.008)	(0.015)	(0.015)	(0.019)
Queue/10 × Green	-0.014	-0.042	-0.006*	-0.014*	0.001	0.003
	(0.012)	(0.031)	(0.003)	(0.008)	(0.003)	(0.007)
Queue/10 × Orange	0.010	-0.004	0.003	-0.006	0.008*	0.002
	(0.012)	(0.026)	(0.004)	(0.009)	(0.004)	(0.011)
Queue/10 × Red	-0.006	0.217	-0.025**	-0.032	0.016	-0.034
	(0.060)	(0.137)	(0.012)	(0.022)	(0.029)	(0.066)
Blue	-1.207***	-1.039***	-0.459***	-0.433***	-0.194***	-0.236***
	(0.071)	(0.127)	(0.011)	(0.020)	(0.022)	(0.020)
Green	-0.855***	-0.822***	-0.226***	-0.218***	-0.161***	-0.166***
	(0.031)	(0.044)	(0.008)	(0.012)	(0.007)	(0.011)
Orange	0.642***	0.660***	0.176***	0.187***	0.157***	0.165***
	(0.021)	(0.043)	(0.008)	(0.013)	(0.007)	(0.015)
Red	-0.360***	-0.681***	-0.010	0.001	0.444***	0.513***
	(0.127)	(0.216)	(0.027)	(0.036)	(0.047)	(0.089)
Female	0.070***	0.066***	-0.008	-0.010	-0.027***	-0.027***
	(0.018)	(0.017)	(0.007)	(0.006)	(0.006)	(0.006)
Age	0.015***	0.015***	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	-0.001**	0.001***	0.001***	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	276061	276061	276061	276061	73609	73609

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results for the interactions between queues and triage colors. For the LOS regressions, results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. For the OOP regressions, the extensive margin uses a total of 276,061 arrivals, from Jan 2011-Oct 2012. The intensive margin uses a total of 73,609 arrivals, from Jan 2012-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level. The benchmark triage color is yellow. First stage SW F-stats: 52.30, 108.29, 461.17, 582.88 and 778.48 (columns 2 and 4); 47.13, 345.63, 218.79, 279.36 and 262.1 (column 6).

other.

Regarding the impact on the out-of-pocket, Table 1.5 shows that the overall negative impact of the queues on the likelihood of being sent for further care is stronger for the low urgency visits, "blue" and "green". The remaining higher urgency visits have a smaller negative impact (in absolute value) that cannot be statistically distinguished from each other. As far as the intensive margin is concerned, none of the color-specific effects are statistically different from zero, suggesting no effect on this margin.

These combined results suggest that doctors ration their time efficiently, in the sense that an exogenous shock to ED visits will lead them to decrease more the time with patients whose condition is not life threatening and, therefore, less likely to suffer from negative health outcomes.

1.4.4 Discharge Destination

In Table 1.6 we run linear probability models on selected discharge destinations. In a first stage we evaluate being admitted to inpatient care against being discharged. In a second stage we compare being sent home against being discharged to primary care services. [14]

Results show that an additional arrival to the ED decreases the probability of a patient being admitted to inpatient care by 0.07 percentage points when compared to being discharged. This corresponds to a 0.61% decrease in the probability of being admitted, at the average probability of admission of 11.4%. Conditional on being discharged, there is no statistical difference between being sent to primary care or home. An increase in the queue increases the likelihood of being sent to primary care or home in similar magnitudes. These findings suggest that doctors react to peaks in arrivals to the ED by sending patients away from hospitals, sending them both to primary care facilities or their homes.

level, as in Figure A.3, only "blue" and "red" are statistically distinguishable.

[14] We choose to model the probabilities of the main discharge destinations sequentially because multinomial models do not easily accommodate the high-level fixed effects that our data exploit for identification. We have tried both a multinomial logit and poisson estimation with high dimensional fixed effects, but did not obtain convergence.

Table 1.6: Discharge destination

Dependent variable:	1 if admitted to inpatient care		1 if primary care | discharged	
	OLS	IV	OLS	IV
Queue/10	-0.007***	-0.042***	-0.026	0.019
	(0.002)	(0.015)	(0.019)	(0.027)
Blue	-0.143***	-0.147***	-0.057***	-0.052***
	(0.004)	(0.006)	(0.010)	(0.009)
Green	-0.112***	-0.114***	-0.092***	-0.089***
	(0.002)	(0.002)	(0.003)	(0.003)
Orange	0.253***	0.252***	0.092***	0.093***
	(0.007)	(0.006)	(0.009)	(0.008)
Red	0.407***	0.407***	-0.210***	-0.211***
	(0.019)	(0.018)	(0.026)	(0.026)
Female	-0.035***	-0.035***	0.103***	0.104***
	(0.004)	(0.004)	(0.006)	(0.006)
Age	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*Age	0.000	0.000	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	222338	222338	197219	197219

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The table shows the results of an LPM on discharge destinations. Results use data from the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level. First stage SW F-stats: 33.91 (column 2); 32.11 (column 4).

1.4.5 Waiting Time

Our dataset does not allow us to analyze the impact of doctors' behavior on patient's health outcomes. Nevertheless, many studies (summarized in a meta-analysis by Hoot and Aronsky 2008) show that ED crowding out is correlated with adverse patient health outcomes, such as mortality. For this reason, we repeat the analysis using waiting times as the outcome variable, in Table 1.7.

The first column tells that one extra person waiting in the ED at the time of check in increases waiting time by 1.67%, on average. At an average queue at check-in of 19.4 patients, the estimates point at an elasticity of waiting time with respect to the queue of 0.32. The second column shows that this effect is true only if the marginal patient shares the same urgency classification, when the effect is stronger and statistically distinguishable from zero. The coefficient for the impact of the queue of different urgency is estimated at a value very close to zero. This result together with those of the length of stay analysis indicate that the arrival of an extra patient of different urgency level does not affect waiting times but may impact their length of stay, after being called to see a doctor. The last column displays the heterogeneity results, which show that the positive effect of queues on waiting times is decreasing with the level of urgency. The result is to be expected since more severe conditions such as heart attacks cannot wait for treatment. In fact, only the two least severe colors, blue and green, present a statistically significant and positive coefficient, while the remaining three cannot be statistically distinguished from zero.

Overall, we find evidence that queues lead to higher waiting times, but the effect is very heterogeneous across urgency levels.

Table 1.7: Waiting time

Dependent variable: Log(Waiting Time)			
Queue at check-in/10	0.167***		0.073
	(0.052)		(0.045)
Same urgency queue at check-in/10		0.360***	
		(0.066)	
Different urgency queue at check-in/10		-0.020	
		(0.041)	
Queue at check-in/10 $\times$ Blue			0.013*
			(0.007)
Queue at check-in/10 $\times$ Green			0.017***
			(0.004)
Queue at check-in/10 $\times$ Orange			-0.015***
			(0.002)
Queue at check-in/10 $\times$ Red			-0.015***
			(0.005)
Blue	0.962***	1.122***	0.712***
	(0.069)	(0.092)	(0.156)
Green	0.470***	0.199***	0.136**
	(0.027)	(0.023)	(0.058)
Orange	-0.960***	-0.757***	-0.666***
	(0.045)	(0.047)	(0.036)
Red	-1.170***	-0.955***	-0.879***
	(0.050)	(0.052)	(0.077)
Female	0.087**	0.084**	0.087***
	(0.034)	(0.034)	(0.034)
Age	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results for the effects of queues on waiting times, using the number of arrivals in the previous 60 minutes as an instrument. In this regression queues are measured at the time of check in. The first column presents the average queue coefficient, while column 2 separates the effects into equal and different urgency queues. The third column shows the heterogeneous effects by iterating queue with urgency level, using "yellow" as the benchmark color. Results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level (15,801 levels). First stage F-stats: 399.78 (column 1); 486.93 and 493.62 (column 2); 929.61, 53.14, 92.68, 49.43 and 46.61 (column 3).

1.5 Robustness Checks

In this section we run several checks that show the robustness of our estimates, which can be found in the online appendix, Appendix A.2. We start by changing our instrument, using only visits with an urgency level ranging from yellow to red, because these are, arguably, the ones that are exogenous to consumer choice (Table A.4). We find that the coefficient of the impact of the queue on the Log(LOS) is of similar magnitude (-0.277 instead of our benchmark of -0.286), and the result that the extensive margin drives the OOP results holds true. We also change the time interval of the instrument, by using arrivals in the previous 15, 30, and 90 minutes (Table A.5). This exercise shows that the F-stat of the first stage and the coefficient estimates decrease as we widen the time interval of the instrument. The elasticities at the mean queue size (14.6) range between -0.38 and -0.5.

We also separate our regression between day and night effects, as the capacity constraints might be very different between the two time periods, and find that our estimates are driven mostly by day-period patients (Table A.6). Finally, we run a model specification that controls for the age of patients in a flexible manner, given that age and health status might not be related linearly (Figure A.4). Results show a monotonic impact of age on LOS and OOP, and the queue coefficient is robust to this change.

1.6 Conclusion

Physicians working in hospital Emergency Departments are often faced with surges in demand. In such circumstances they need to readily adapt, optimizing the resources spent with each patient. In practice this means that whenever the ED is (unpredictably) crowded, doctors may redefine their allocation of resources by decreasing the amount of tests and time spent with each patient.

Using an instrumental variable strategy along with a fixed-effects model, we provide causal evidence that physicians treat patients less intensively in periods of spiking demand. The interval between being first seen by the doctor and discharge decreases when more patients are queuing in the waiting room, as does the likelihood of being sent to further diagnosis and treatment after seeing the doctor. Also, patients are more likely to be discharged from the hospital instead of being admitted to inpatient care.

Rationing both time and lab tests is not bad *per se*, and whether it has negative health consequences is an empirical question that depends on the health production function. We do not test the health outcomes coming from variation in arrivals since our data does not allow us to check readmissions or deaths in a reliable way. Nevertheless, our results do suggest that health impacts might be a valuable topic for further analysis, as we find evidence that arrivals are linked to changes in waiting times, known to be correlated with adverse health outcomes (Hoot and Aronsky 2008). This paper focuses, instead, on studying a mechanism through which ED visits and health outcomes can be linked - the physicians' role - and leaves the assessment of its consequences for future research. It is encouraging that we find some evidence that physicians ration efficiently, in the sense that they decrease treatment more intensively for patients without urgent conditions, whose health impacts are more likely to be small.

Overall, our study has important policy implications. Our results show that doctors react to capacity constraints and, therefore, policies aimed at changing the amount of resources available to the hospital need to account for doctors' reactions. For example, hiring additional doctors in order to decrease waiting times might have a lower than expected impact if the infra-marginal doctors spend more time with patients as a consequence.

Part II

Inpatient care

Chapter 2

A bed constraint? Occupancy rates and hospital readmissions

2.1 Introduction

Universal provision of health care is one of the main foundations of European modern societies. However, generating sufficient hospital capacity to guarantee everyone receives ideal treatment at any given time is a sensitive matter. Since demand for health care is unpredictable, the optimal level of capacity must balance potential health consequences of peak-load times against the opportunity cost of resources under idle times. Consequently, it is common for public hospitals to experience unexpected demand fluctuations that lead to acute bed shortages (Bagust, Place and Posnett, 1999; Bain et al., 2010; Green and Nguyen, 2001; Gorunescu, McClean and Millard, 2002; Bekker and Koeleman, 2011). When hospitals are full and resources are insufficient, operational failures are likely to occur (Epstein et al., 2012; Tucker, 2004). Importantly, doctors' decisions regarding admissions and discharges may be affected by the pressure to clear beds and admit patients in urgent need of hospitalization (Blom et al., 2014*b*). For example, physicians may decide to admit only severe patients, or discharge patients prematurely to make room for other patients.

Intuitively, physicians, who decide the quantity of resources to apply based on the characteristics of the patients and service constraints, may have to ration resources, such as their own time, to treat more people (Clark and Olsen, 1994; Cutler, 1993; Whynes, 1996). Time constraints drive doctors to optimize based on a time-quality trade-off (Anand, Paç and Veeraraghavan, 2011; Dugdale, Epstein and Pantilat, 1999). For instance, they may decrease the length of stay during critical periods to make room for more admissions (Bagust, Place and Posnett, 1999; Forster et al., 2003; McCarthy et al., 2009; Thomas and Holloway, 1991). These choices are important, as time spent with patients is related to positive health outcomes (Ogden et al., 2004; Chen, Farwell and Jha, 2009; Silver, 2016).

Occupancy rates have been found to be correlated with negative outcomes in emergency departments, particularly in intensive care units (Chrusch et al., 2009; Gattinoni et al., 2004; Boden et al., 2015; McCusker et al., 2014). They are also related with emergency department length of stay (Bagust, Place and Posnett, 1999; Forster et al., 2003; McCarthy et al., 2009). Finally, the association between occupancy rates and readmission episodes is mixed, switching from positive (Chrusch et al., 2009; Durbin et al., 2007; Blom et al., 2015), to negative or non conclusive (Blom et al., 2014*b*,*a*), depending on the type of hospital or department, and the type of readmission interval chosen (e.g., 30 days vs. 72 hours readmission).

Understanding the effects of high occupancy rates is fundamental to gauge their potential costs and, if needed, appropriately design policies to tackle inefficiencies. For that purpose I assess whether demand-supply mismatches impact the quality of inpatient care. Moreover, I distinguish the effects by age groups, giving special emphasis to the elderly population. I use readmission episodes as a quality measure (Benbassat and Taragin, 2000), thus aiming at defining the effects of occupancy levels (number of patients per number of beds) on the likelihood of readmission episodes in Hospitals of the Portuguese National Health Service (NHS).

In this study I use data on the Portuguese Diagnosis Related Groups (DRG), provided by the Portuguese Central Administration of the Health System (ACSS), containing anonymized individual-level data on all discharges in Portuguese public hospitals. I look at how occupancy rates, and respective variation, in days prior to patients' discharge, are associated with readmission episodes. I leverage my analysis on a fixed effects model to exploit a fine variation in weekly occupancy rates taking place in each hospital, while accounting for the DRG associated with each discharge and its level of severity. Additionally, I perform a cross-age comparison to determine if the impacts are constant across different age groups. Finally, I address the relation between occupancy rates and early discharges, using exogenous thresholds of length of stay. My analysis provides important insights regarding the consequences of occupancy rates on the provision of inpatient care. At the same time, it addresses potential inequalities of care provision on different age groups.

Results show that occupancy rates faced by patients in one and two days prior to discharge cause the likelihood of future readmissions episodes to go up. The same happens when occupancy variation is used instead of its level. Results are driven by the older groups, which tells that the phenomenon affects especially the frailer population. The length of stay threshold analysis suggests higher probability of readmission for patients discharged below the expected length of stay for a given level of severity of a given DRG, indicating early discharges as the main cause. Again, the result is driven by the older age groups in the sample.

This paper is close to Blom et al. 2015, as I also assess the effects of occupancy on 30 days readmissions. I contribute to the literature by defining this relation for Portuguese public hospitals, where physicians do not have any performance-related financial incentive, using a finely detailed fixed effects model. Also, I go deeper in the analysis of potential early discharges, providing further explanation to the main findings. Finally, I study health care access inequalities by identifying the most vulnerable age groups. This information is useful to policy makers, as for instance, they can conclude that when occupancy rates increase, physicians need to be especially careful when discharging older patients with below expected length of stay.

The paper is presented in the following order: Section 2.2 provides a brief institutional context for the inpatient care services in the Portuguese NHS. Section 2.3 presents and explains the data and the relevant variables. Section 2.4 debates the methodology adopted. Section 2.5 exhibits and discusses the results of the effects of occupancy rates on readmission episodes. Section 2.6 discusses the relation between occupancy and length of stay and Section 2.7 provides additional results. Section 2.8 concludes.

2.2 Institutional context

The Portuguese health care system is primarily characterized by a universal National Health Service (NHS)[1], that is regulated at the federal level. A comprehensive description of the Portuguese health care system can be found in Barros, Machado and Simões 2011.

All residents in Portugal are covered by the NHS, irrespective of their socioeconomic, employment or legal status. The access to the public inpatient care service is usually made through the emergency care department for acute conditions, or through general consultations with specialized physicians for non-acute conditions. In urgent cases, individuals are admitted to the nearest hospital, unless the hospital has no specialized department with the ability to treat a specific condition. In that case, the patient is readily transferred to a specialized hospital. In less urgent cases, derived from non-acute care, the patient is normally admitted to the hospital in her area of residence, in which she was being followed by a specialty doctor.

[1]The public sector accounts for about 80% of emergency care and more than 70% of total hospitalizations and surgeries, according to the Portuguese Institute of Statistics (INE).

After being hospitalized, the length of stay of each patient is defined by the doctor responsible for the treatment. Independently of the period of hospitalization, there is no user charge, and therefore the service is considered free.[2]

Hospital budgets are drawn up and allocated by the Ministry of Health through the Central Administration of the Health System (ACSS). Public hospitals are allocated a global budget using DRGs to adjust the prospective budget for case-mix and other hospital specificities. This need of collecting individual-level data for DRG grouping purposes has significantly improved the hospitals' information system.[3]

NHS doctors are salaried government employees with a fixed salary that depends on the category and duration of service. There is no pay for performance schemes in inpatient care, which means doctors have no financial incentives. Still, it is possible for doctors to work both in the public and private sector. Currently, there are three employment levels for doctors: full-time, but not exclusive, full-time with exclusive NHS employment, and part-time. Full-time practitioners need to work 40 hours in a week.

The public system owns 48 institutions that are separated in 4 different groups: hospital centers, general hospitals, specialized hospitals and local units of health (ULS). The hospital centers are clusters of hospitals with functional connections under the same administration. General hospitals are stand-alone hospitals with their own management. They provide mostly the same services as the hospitals in the hospital centers. Specialized hospitals focus on a limited number of services, usually associated with specific types of conditions (e.g. oncology, psychiatry or ophthalmology). Finally, local units of health integrate several services of the health system such as hospital care and primary care.

Among the 48 institutions, there are currently eight medical schools in Portugal, two in Lisbon, two in Oporto and one in Coimbra, Braga, Covilhã and Algarve).

[2] Payments to the NHS are made indirectly, through taxes.

[3] That is the data I use in this paper.

2.3 The Data

2.3.1 General overview

The data was provided by the Portuguese Central Administration of the Health System (ACSS) and contains yearly collected DRGs' (Diagnosis Related Groups) data for all admissions in the Portuguese public health institutions. Importantly, each patient is associated with an identification code, thus allowing me to find readmission episodes. I collected a complementary data-set to account for the monthly number of beds in each hospital, available in the institutional site of ACSS.

The DRG data-set contains information that identifies hospitals, date of admission, type of admission[4], date of discharge, type of discharge[5], transfer motives [6], and diagnoses. It also includes age, gender and place of residence of the patient.

The initial sample is comprised of all individuals hospitalized in all hospital centers of the Portuguese NHS and whose discharge date lies between the 1^{st} of January of 2014 and the 31^{th} of December of 2016.

Outpatient care is dropped from the sample because admissions are scheduled in advance, meaning demand is predictable. Also, they mostly compete for different resources within the hospital. I also exclude December of every year because of the impossibility of counting readmission episodes generated during that month, due to data structure[7]. Finally, I keep only the seven biggest Major Categories of Diagnosis in the sub-sample, after excluding childbirth related episodes. The final sample contains 41 hospital centers and roughly 1.7 million hospitalizations.

[4]Planned, unplanned, hospitalization or ambulatory.
[5]Ordered, or not, by a physician
[6]Follow up, lack of equipment, long-term care, etc...
[7]Patient identifier does not transition across years

In the scope of this truncated data-set, the sample individuals average 63 years of age, with a slight discrepancy in gender in favor of men (52.12%).

The seven MDC (Major Diagnosis Categories) included in the sub-sample are the respiratory system (19.01%), circulatory system (17.96%), musculoskeletal system (17.06%), digestive system (15.12%), nervous system (11.22%), urinary system (10.81%) and hepatobiliary system (8.83%).

In the Portuguese NHS, patients are not allowed, at the time of the data, to decide in which hospital they want to be treated, being automatically sent to be hospitalized in either their hospital of residency or the nearest hospital, in case of emergency.

2.3.2 Readmission episodes:

I built the readmission variable using individuals' identification codes. The numerical identification codes associated with each patient allow me to trace the people treated in the NHS's hospitals, and therefore to count re-hospitalizations. Readmissions occur when patients are re-hospitalized following a previous hospitalization (index admission). I use two approaches when counting readmissions, and both with a 30 days threshold[8]. Despite some debate, this definition has been widely used in the literature (Benbassat and Taragin, 2000; Boulding et al., 2011; Blom et al., 2015). First, I adapt the Medicare definition, considering all admissions, independently of the cause, to constitute readmissions if they happen within 30 days after discharge on a first admission. This method will be called the standard approach. Second, I count readmission episodes when the time condition is met (30 days) and the readmission's MDC is the same as the index admission. In this scenario, readmissions are more likely to be related with the initial admission. This method will be called the conservative approach. Both methods have into account cross hospital admissions as

[8] Admissions separated by more than one month are unlikely to be related.

long as they are unplanned.[9]

Readmissions are counted at the index admission, so it is possible to study the association between occupancy rates and future readmission episodes.[10] Moreover, readmission episodes are only counted if discharges associated with the index admissions were ordered by a physician and the readmissions are unplanned[11].

In the conservative approach, readmissions amount to 3.91% of the sample, while in the standard approach they represent 7.75% of the total observations. The respiratory system (13.5%), the urinary system (12%) and the hepatobiliary system (11%) are the MCDs with higher readmission rates, in the standard approach. In the conservative approach, the highest readmission rates occur in the musculoskeletal system (12%), respiratory system (8%), and the hepatobiliary system (7%).

For hospitals, using the standard approach, the data shows a maximum of 10.5% average readmission rate in a hospital center located in the middle interior of the country. The minimum value, 4%, belongs to a small hospital in an occidental part of the northern region of Portugal. The median hospital in terms of readmission rates presents a value of 6.80% and is located in the city of Lisbon.

Importantly, the month of December is excluded because the identification code is lost across years. As an example, if a patient were to be discharge in December 31th and then readmitted in January, the identification code would be different, which prevents the identification of the patient and consequently counting the readmission episode.

[9] Cross hospital admissions are accounted as potential readmissions if they are not transfers.

[10] The existence of a time gap avoids the possibility of reverse causality.

[11] The data provides a code stating whether an admission was urgent or previously scheduled.

2.3.3 Occupancy rates:

The information concerning the dates of admission and discharge of the patients allow me to build variables of occupancy in each hospital of the Portuguese NHS. The ratios of occupancy are calculated by counting the number of patients hospitalized in each day of the year dividing by the number of beds[12], in the respective hospitals. Figure 2.1 illustrates the evolution of occupancy rates along the three years contemplated by this study.

Figure 2.1: Monthly Occupancy Rates

Note: This figure shows the evolution of the average occupancy rates that patients' experienced in their last day of hospitalization in all months of the years in the analysis (2014, 2015 and 2016).

The graph displays what seems to be a negative trend within each year for all years in analysis.

[12]I use the variable in its ratio format, which means that 0.8 corresponds to 80% occupancy rate.

August is consistently the lowest occupancy rate month because it is traditionally a month of vacations.

I measure occupancy just before the discharge date, that is, in the last day of hospitalization, or alternatively, in the days just prior to discharge. The measures aim at capturing potential effects where sudden peaks in occupancy cause early discharges.

In the data, the average rate of occupancy at discharge is 0.9731 (97.31%). This number indicates that, on average, Portuguese patients are likely to find very crowded hospitals during their hospitalization. The results, however are an overestimation of the real value[13]. This happens because the measure counts individuals that were in the hospital in a given day. Some patients may not overlap their stay, but are still counted as such.[14] If a patient is discharged in the morning and another patient is admitted during the afternoon, then they faced different occupancy rates, even though they are being counted together. Therefore, there is a problem of double counting that inflates the values. Still, since the direction of the measurement error is known, the overestimation does not compromise the qualitative conclusions. Moreover, the analysis includes a turnover rate to account for that problem. The turnover rate is the sum of admissions and discharges on the day of interest, and it is expected to capture how busy the hospital was, filtering the occupancy variable from the double counting issue.

Finally, I also compute the variation in occupancy which is I use throughout this study to complement the main analysis. The variation in the occupancy rate is computed by subtracting the number of admissions and discharges in a given day, in a each hospital.

[13]The occupancy rates should be approximately 5 percentage points lower, according to data in ACSS's institutional site.

[14]Not all years and hospitals in the data have the hour of admission and discharge. As such, I cannot fully correct the measurement error.

2.4 Methodology

Regression analysis is used to assess the potential relation between the negative outcomes and the capacity variables. My main hypothesis is that doctors force discharges in periods of greater occupancy, therefore increasing the likelihood of early discharging and consequently the probability of readmission. The general model I use to test this theory is written in the following way:

$$y_{iht} = \alpha + \beta Occupancy_{iht} + \gamma X_{ith} + \theta DRG_i \times Severity_i + \tau_y \times \tau_m \times \tau_w \times \tau_h + \varepsilon_{iht} \quad (2.1)$$

The dependent variable, y_{iht}, is a binary variable stating whether discharge i, in hospital h, with discharge at time t (day), generated a readmission. The main explanatory variable, $Occupancy_{iht}$, contains occupancy-rate related variables for discharge i in hospital h at day t. I conduct experiments using the occupancy variable as either occupancy at discharge, occupancy one day before discharge, occupancy two days before discharge and occupancy three days before discharge[15]. I also use occupancy variation rates in a complementary analysis.

The interaction $\tau_y \times \tau_m \times \tau_w \times \tau_h$ provides weekly fixed effects interacted with hospital fixed effects, meaning that the model studies a very fine variation in readmission episodes in a given week of the year in a specific hospital. Variable DRG_{it} stands for the respective diagnosis related groups' fixed effects. The interaction between week and hospital accounts for seasonal variation and epidemics, and hospital characteristics. DRGs' fixed effects are used to control for cross diagnosis severity. They are interacted with four levels of severity[16] to provide a finer control for the patient's condition.

[15]The lags stop at three days before discharge because it is unlikely that physicians decide discharges sooner than that.

[16]1-low, 2-medium, 3-high, 4-extreme.

Vector X_{ith} contains a gender variable, age groups, number of procedures and a daily turnover rate. The number of procedures aims at further accounting for severity. The daily turnover rate is a variable with the sum of the number of admissions and discharges in the day in which patient i is discharged, as a proportion of total patients that where in the hospital that day. Variable $gender_i$ is a dummy taking value 0 if the individual is female. Variable age_{it} is composed of age intervals of 5 years from 0 to 10 and 10 years from the age of 10 to 90. Individuals older than 90 are grouped in the last interval. Age is measured in groups since the effect is likely to be non-linear.

The analysis starts with evaluating occupancy at discharge day, in line with Blom et al. 2015 and then continues testing the lags of the variable. The latter assumes that physicians decide discharges beforehand, and thus the results are likely to be more prevalent in the days prior to discharge. The model is estimated using a Linear Probability Model (LPM) to more easily accommodate the high degree of fixed effects. The out of bound predictions in the main estimations amount to less than 3%.

2.5 Results

2.5.1 Occupancy in the day of discharge

Table 1 shows the average marginal effects of occupancy rates at the day of discharge (following Blom et al. 2015) on the probability of readmission.

In the two approaches, results show that being discharged in a day with higher occupancy rates is not affecting readmissions. The estimates are small and non-statistically significant.

However, the daily turnover rate has a positive relation with the probability of readmission. A 10

Table 2.1: Main Estimates on the Probability of Readmission

Dependent variable: Readmission				
	Standard	Conservative	Standard	Conservative
Occupancy at Discharge	0.008	0.008		
	(0.008)	(0.005)		
Turnover Rate	0.053***	0.025***		
	(0.003)	(0.002)		
Occupancy Rate Variation			-0.107***	-0.059***
			(0.006)	(0.005)
Male	0.008***	0.006***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Procedures	0.002***	0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Age [5-10]	-0.007**	-0.000	-0.007**	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Age [10-20]	-0.007**	-0.001	-0.007**	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Age [20-30]	-0.011***	-0.005**	-0.010***	-0.005**
	(0.002)	(0.001)	(0.002)	(0.001)
Age [30-40]	-0.015***	-0.007***	-0.014***	-0.007***
	(0.002)	(0.001)	(0.002)	(0.001)
Age [40-50]	-0.014***	-0.006***	-0.013***	-0.006***
	(0.002)	(0.001)	(0.002)	(0.001)
Age [50-60]	-0.009***	-0.004**	-0.008***	-0.003*
	(0.002)	(0.001)	(0.002)	(0.001)
Age [60-70]	-0.002	-0.002	-0.001	-0.002
	(0.002)	(0.001)	(0.002)	(0.001)
Age [70-80]	0.005**	-0.001	0.006**	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
Age [80-90]	0.017***	0.001	0.017***	0.001
	(0.002)	(0.001)	(0.002)	(0.001)
Age [90+]	0.021***	-0.001	0.021***	-0.000
	(0.002)	(0.001)	(0.002)	(0.001)
R^2	0.057	0.036	0.057	0.036

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of occupancy rates at discharge on the probability of a readmission episode. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels) and DRG-Severity fixed effects (595).

percentage point increase in the turnover rate would lead to an increase in the probability of readmission of 0.5 and 0.25 percentage points, in the standard and conservative approaches, respectively. Both explain roughly 5% of the respective readmission rates (at rates 8.94% and 4.49%). The results suggest that busier days are positively related to the likelihood of readmission episodes.

The occupancy rate variation, in the two last columns, states that patients discharged in a day when hospital occupancy rates have increased have a lower probability of readmission. One possible interpretation is that days with very high number discharges, and consequently negative variation in occupancy, are associated with higher probability of readmission.

The gender marginal effects indicate that men are more likely to be readmitted than women. Men have a higher probability of readmission between 0.6 and 0.8 percentage points, on average, depending on the methodology. The result is robust across all specifications of the model presented in this manuscript.

The number of procedures impacts readmission probability positively, with individuals that are subject to an extra procedure experiencing an increase in the probability of being readmitted of 0.2 percentage points, in the standard approach. Naturally, more severe cases are associated with both a higher number of procedures and a higher readmission probability.

Finally, the age interval coefficients in this LPM shows that the probability of readmission increases with age. As people get older, they become frailer, taking longer to recover. This result holds in the standard approach but not in the conservative approach.

2.5.2 Occupancy before discharge

In this section, I test a different explanatory variable. Instead of occupancy in the day of discharge (more common in the literature), I look at the effect of the occupancy variable before discharge. The hypothesis underlying this test is that doctors decide about discharges in advance. Thus, occupancy variables before discharge should be more important on explaining discharging decisions than occupancy variables in the discharge day itself.

Table 2.2 displays the results using occupancy one day before discharge as the explanatory variable. The results using occupancy rate two and three days before discharge are shown in appendix, in Table B.1 and Table B.2.

Table 2.2: Main Estimates on the Probability of Readmission

Dependent variable: Readmission				
	Standard	Conservative	Standard	Conservative
Occupancy One Day Before Discharge	0.070***	0.036***		
	(0.008)	(0.006)		
Turnover Rate t-1	-0.006*	-0.002		
	(0.003)	(0.002)		
Occupancy Rate Variation t-1			0.041***	0.015**
			(0.007)	(0.004)
Male	0.008***	0.006***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Procedures	0.002***	0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.056	0.036	0.056	0.036

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of occupancy rates one day before discharge on the probability of a readmission episode. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

Under this alternative, the coefficient on occupancy is statistically significant for both the standard and conservative approaches. According to the estimates, a 10 percentage point increase in the occupancy rate one day before discharge leads to a 0.7 percentage point increase in the likelihood of readmission, in the standard approach. In the conservative approach, the estimates tell that a 10 percentage point increase in the occupancy rate one day before discharge increases thereadmission probability by 0.36 percentage points.

In terms of magnitude, a back of the envelope example shows that increasing occupancy by 10 percentage points, at marginal effects of 0.07 or 0.036, for 1,303,129 individuals in this sample, with an average length of stay of 9.4 days and a hospitalization average cost per day of 880 euros[17] yields an extra cost to the Portuguese NHS of 38.8 or 75.5 million euros in three years[18], depending on the approach chosen. The example is crude and purely illustrative. A more thorough financial analysis would require the exact cost of each episode, which is not available.

The two last columns of Table 2.2, report the coefficients of the lagged variation in occupancy for the two approaches. Contrary to the results found in Table 2.1, a higher variation in occupancy positively affects readmissions. When variation in occupancy increases by 10 percentage points the readmission probability increases by 0.15 and 0.41 percentage points in the conservative and standard approach, respectively.

These results indicate that there is a positive relation between occupancy rates and readmissions, but only prior to discharge. The findings suggest that doctors take into account the state of occupancy of the hospital when taking discharging decisions which should occur at least one day before the effective discharge. Physicians take into account not only the absolute value of occupancy in that day, but also the observed variation in occupancy. This is consistent with my hypothesis that they force discharges when hospitals are busier to make room for new patients, which increases the probability of patient readmission.

[17]The value was computed by the Portuguese Health Regulator.

[18]The summed up value of 75.5 million euros in the standard approach for the three years in the analysis accounts for roughly 0.05% of the 2016 Portuguese GDP.

Table B.1 and Table B.2, in the appendix, present the same estimation for occupancy two and three days before discharge. The results for occupancy are smaller when measured two days before discharge (but are still positive and statistically significant).When looking at three days before discharge, the coefficient on occupancy is negative and not significant. In terms of occupancy variation, the two-days lag shows higher coefficients than the one-day lag. The three-days lag coefficients are also positive and statistically significant, but lower than the other two. Summarizing, early discharge decisions depend mostly on how occupancy rates are evolving two days before discharge (occupancy variation) and how they fare, in absolute terms, one day before discharge (occupancy rate).

2.5.3 Occupancy rates by intervals

So far, the effect of occupancy rates on the probability of readmission has been studied assuming a continuous effect. However, it is likely that occupancy rates only matter past a critical threshold, when medical time becomes a scarce resource. For example, occupancy rates below 80% should not be problematic (Bagust, Place and Posnett, 1999), as the hospital is relatively empty. Thus comparing 70% occupancy with 80% occupancy rates should not show significant effects on readmission probabilities. But comparing 90% occupancy rate and 100% occupancy rate is likely to provide meaningful results. In that regard, this section evaluates the hypothesis that occupancy rates affect readmission episodes in a non-linear fashion. So, instead of using continuous variables of occupancy, Table 2.3 presents regression results by intervals.

Results in Table 2.3 show that occupancy intervals evaluated one day before discharge have a weakly increasing effect on readmission rates, with the higher interval presenting the highest coefficients in all specifications. The interval of 110% and above has a coefficient of 0.014 in the standard approach and 0.008 in the conservative approach. This implies that a patient discharged under those circumstances is 1.4 or 0.8 percentage points more likely to be readmitted when compared to a patient whose occupancy one day before discharge was below 85%. These values are

Table 2.3: Main Estimates on the Probability of Readmission

Dependent variable: Readmission				
	Standard	Conservative	Standard	Conservative
Occupancy Rate $>$ 110% t-1	0.014***	0.008***	0.019***	0.010***
	(0.002)	(0.002)	(0.002)	(0.001)
Occupancy Rate [105%, 110%] t-1	0.009***	0.006***	0.012***	0.007***
	(0.002)	(0.001)	(0.001)	(0.001)
Occupancy Rate [100%, 105%] t-1	0.007***	0.005***	0.010***	0.006***
	(0.002)	(0.001)	(0.001)	(0.001)
Occupancy Rate [95%, 100%] t-1	0.006***	0.004***	0.007***	0.004***
	(0.002)	(0.001)	(0.001)	(0.001)
Occupancy Rate [90%, 95%] t-1	0.003	0.003**	0.003*	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
Occupancy Rate [85%, 90%] t-1	0.006***	0.003*	0.007***	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)
Turnover Rate t-1	0.001	0.001		
	(0.002)	(0.002)		
Occupancy Rate Variation t-1			0.068***	0.029***
			(0.007)	(0.005)
Male	0.008***	0.006***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Procedures	0.002***	0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.056	0.036	0.057	0.036

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of different occupancy intervals on the probability of a readmission episode. The interval of 0% to 85% is the base interval. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

respectively 15.7% and 17.8% of total readmission probability.

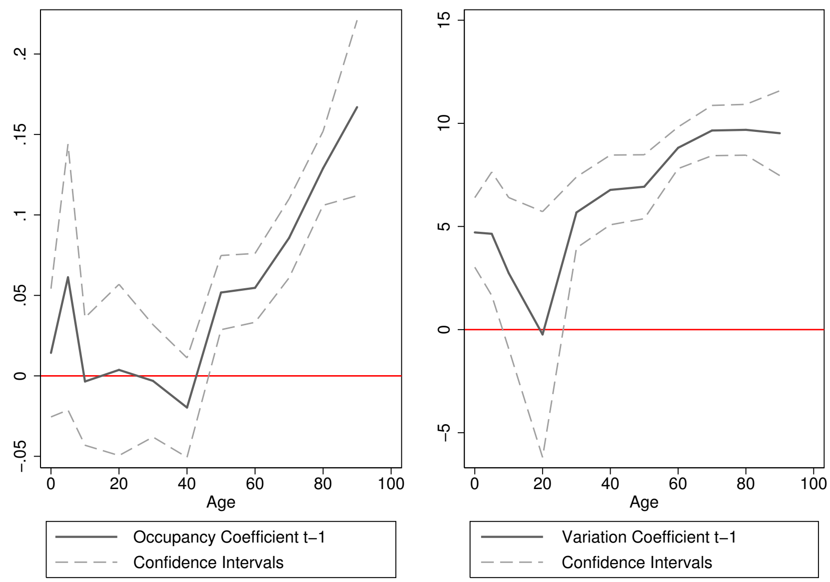
The results for occupancy rate variation are shown in the appendix, in Appendix B.1, in Figure B.1 and Figure B.2. They show that occupancy rate variation one day before discharge is only relevant in predicting readmissions if occupancy rate is high enough (greater than 85%).

2.5.4 Occupancy and readmission by age groups

My second objective is to determine if there is treatment inequality across age. To find out, I perform individual regressions per age group. The model is estimated in a similar fashion to the one described by Equation (2.1). The explanatory variable is occupancy one day before discharge. A complementary approach looks at occupancy variation one day before discharge. The coefficients on the occupancy and variation variables for each age group are depicted in Figure 2.2, using the standard approach. The figure for the corresponding conservative approach is in Figure B.3, in the appendix.

The figure shows that readmission episodes associated with occupancy rates at discharge may be attributed to older patients. In fact, only the age groups over 50 years of age present positive and statistically significant coefficients. All the other groups have statistically non-significant coefficients, and some are negative. The variation coefficients also present a similar trend. After 30 years, they are increasing with age. The curve however is not as steep as the occupancy one. The combined results provide evidence that there is inequality in the outcomes across age groups. Older people lose more by being discharged when occupancy rates are high and increasing, as their chance of readmission is higher.

Figure 2.2: Occupancy Coefficients by Age



Note: The figure shows the effects of occupancy rate and occupancy variation one day before patient's discharge on the likelihood of readmission by age, using the standard approach.

2.6 Length of stay

2.6.1 Length of Stay and Occupancy

The hypothesis established in this paper for the effect of high occupancy rates on readmissions says that sudden demand spikes can lead to early discharges which in turn lead to worse health outcomes. Therefore, a complete analysis must evaluate the relation between occupancy rates and patients' length of stay, and the consequent effect on readmissions.

First, I run the model in Equation (2.1) with length of stay (measured in days) as the dependent variable. Table 2.4 shows that occupancy variables are related with higher length of stay. Occupancy one day before discharge has positive coefficients, suggesting that patients discharged when capacity is limited tend to experience larger periods of hospitalization. In column (1), a 10 percentage point increase in occupancy one day before discharge is associated with roughly half a day

Table 2.4: Main Estimates on the Length of Stay

Dependent variable: Length of Stay	(1)	(2)
Occupancy One Day Before Discharge	4.205***	
	(0.317)	
Turnover Rate t-1	-2.063***	
	(0.105)	
Occupancy Rate Variation t-1		8.079***
		(0.246)
Male	0.158***	0.158***
	(0.020)	(0.020)
Number of Procedures	1.000***	0.999***
	(0.007)	(0.007)
Age [5-10]	-0.473***	-0.478***
	(0.059)	(0.060)
Age [10-20]	-0.420***	-0.422***
	(0.073)	(0.073)
Age [20-30]	-0.438***	-0.442***
	(0.096)	(0.097)
Age [30-40]	-0.263***	-0.268***
	(0.072)	(0.072)
Age [40-50]	0.028	0.023
	(0.070)	(0.070)
Age [50-60]	0.388***	0.384***
	(0.066)	(0.067)
Age [60-70]	0.694***	0.687***
	(0.064)	(0.065)
Age [70-80]	0.799***	0.793***
	(0.067)	(0.067)
Age [80-90]	0.772***	0.766***
	(0.072)	(0.073)
Age [90+]	0.027	0.027
	(0.078)	(0.078)
R^2	0.219	0.219

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of occupancy rates one day before discharge on patient's length of stay (in days). Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels) and DRG-Severity fixed effects (595).

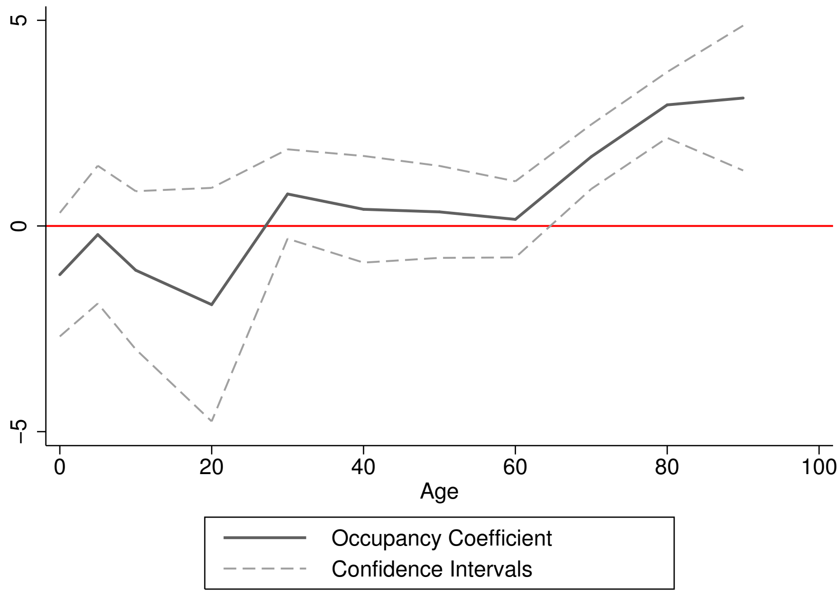
more of length of stay. Occupancy variation one day before discharge follows the same trend. In column (2), an increase in 10 percentage points in occupancy rate variation leads to almost one full day of additional length of stay.

The results may seem counter-intuitive, as the hypothesis of early discharges suggests that it would be more natural to observe shorter length of stay for patients discharged when occupancy rates are high. However, it is possible that results are a product of reverse causality. Intuitively, physicians may try to optimize discharges by sending home patients who have been in the hospital for a longer period. Alternatively, patients who stayed in the hospital for longer may have suffered from more severe conditions even withing their reported DRG and lever of severity. In a situation where doctors have to make room for new admissions, they could choose to discharge patients with longer length of stay, conditional on their severity level. The result is consistent with Martins and Filipe 2020, who say that doctors decrease length of stay for lower urgency patients in emergency care, when the emergency department is more crowded. In an inpatient care setup, the same type of decision making may be expected.

As before, the analysis can be decomposed into age groups, indicating who are the people being caught up in the mechanism of discharges due to high occupancy rates. Results are presented in Figure 2.3. The figure shows that older people discharged at higher occupancy rates one day prior to discharge are more likely to have higher length of stay. It is possible that doctors, when faced with higher occupancy rates, decide to discharge patients who have been hospitalized for longer. Some of these patients may be older individuals, who, in general, need more time to get better for a given illness and therefore have stayed in the hospital for longer.

A joint analysis of this last figure and Figure 2.2, from Section 2.5.4, seems to imply that because the elderly stay in the hospital for longer they are more prone to being discharged when occupancy rates are higher and thus, they have a higher chance of readmission. However, the methodology so far has not presented any evidence that the patients with higher length of stay are the ones more likely to return to the hospital. Also, it is hard to say if readmissions are caused by early

Figure 2.3: Length of Stay by Age



Note: This figure shows the effects of occupancy rate one day before patient's discharge on the length of stay.

discharges, as length of stay is not enough to determine if the patient left the hospital earlier than recommended. The next section tackles this issue using exogenous thresholds of length of stay for DRG and level of severity.

2.6.2 Length of stay by thresholds

In this subsection, I aim at determining if the results in previous sections are a product of early discharges. Determining whether a patient was discharged sooner than recommended is a complicated affair. Even physicians may diverge in opinion if presented specific cases of potential early discharges. That said, I opted for a very objective approach, relying on exogenous thresholds to evaluate the effect of occupancy variables on patients within different categories of length of stay. I use the table, defined by the Portuguese law that prices inpatient episodes [19], which depends on some intervals of length of stay, for each DRG and severity level. The intervals are divided in 5,

[19]Available in portaria Portaria n.º 254/2018.

sequentially called minimum length of stay, below expected length of stay, above expected length of stay, superior length of stay and maximum length of stay. They are set ex-ante by the Portuguese authorities, which makes them independent of the length of stay observed in the data. Interacting each interval with the occupancy variables tells, on average, how patients with observed higher occupancy rates (or higher occupancy rate variation) one day before discharge had their probability of being readmitted change depending on their relative level of length of stay. Table 2.4 shows the results of occupancy variables interacted with thresholds of length of stay. The regression follows Equation (2.1) of section Section 2.4, only substituting the occupancy variable by the aforementioned interaction variables.

Results show three important trends. First, patients discharged below expected length of stay, for their DRG and respective severity level, see their likelihood of readmission increase if occupancy variables increase. Specifically, for that group, an increase of 10 percentage points on occupancy rates one day before discharge implies an increase in readmission probability of 0.2 and 0.35 percentage points, in the conservative and standard approaches, respectively. The coefficients are higher for the variation variable, with an increase of 10 percentage points in occupancy variation one day before discharge leading to an increase in the probability of readmission of 1.5 and 1 percentage points, in the conservative approach and standard approach, respectively. Second, positive changes in occupancy variables one day before discharge negatively affect the readmission probability of patients whose length of stay is above expected. Third, increases in occupancy variables increase the probability of readmission for patients with maximum length of stay. The coefficients are not as large as the above expected length of stay group, but they are significant in all specifications.

The interaction between occupancy variables and below expected length of stay indicate that higher occupancy variables do indeed cause early discharges, since this is the group suffering the most in terms of readmission likelihood. The interaction between occupancy variables and above ex-

Table 2.5: Main Estimates on the Length of Stay by Thresholds

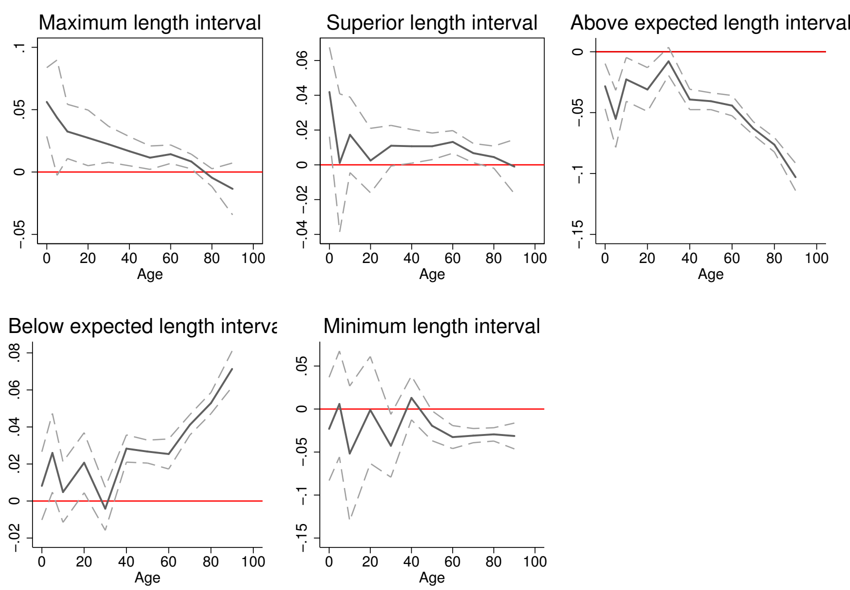
Dependent variable: Readmission				
	Standard	Conservative	Standard	Conservative
Occupancy t-1 & Max Limit Length	0.010***	0.004***		
	(0.002)	(0.001)		
Occupancy t-1 & Superior Limit Length	0.008***	0.002		
	(0.001)	(0.001)		
Occupancy t-1 & Above Expected Length	-0.055***	-0.029***		
	(0.001)	(0.001)		
Occupancy t-1& Below Expected Length	0.035***	0.020***		
	(0.001)	(0.001)		
Occupancy t-1& Min Limit Length	-0.037***	-0.016***		
	(0.003)	(0.002)		
Turnover Rate t-1	0.021***	0.011***		
	(0.002)	(0.001)		
Variation t-1& Max Limit Length			0.151**	0.069*
			(0.049)	(0.029)
Variation t-1& Superior Limit Length			0.046	0.056*
			(0.038)	(0.025)
Variation t-1& Above Expected Length			-0.143***	-0.106***
			(0.038)	(0.026)
Variation t-1& Below Expected Length			0.150***	0.106***
			(0.037)	(0.027)
Variation t-1& Min Limit Length			0.116*	0.081*
			(0.056)	(0.038)
Male	0.008***	0.006***	0.009***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Procedures	0.001***	-0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.059	0.037	0.056	0.036

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of occupancy rates interacted with thresholds of occupancy defined in the scope of the financing of the Portuguese Public Hospitals. The thresholds are defined per DRG and severity level. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels) and DRG-Severity fixed effects (595).

pected length of stay say that in this specific group occupancy rates cause no additional probability of readmission, giving strength to the previous conclusion. Finally, the maximum limit of length of stay is more puzzling, since its interaction with occupancy variables provides a positive coefficient. Some likely scenarios could incorporate phenomena such as social hospitalization, i.e., patients who have nowhere to go and therefore must return to the hospital. Alternatively, it could be due to extreme cases, such as accidents, where the number of complications go beyond what can be expected or incorporated in a single diagnosis related group.

Figure 2.4: Length of Stay Thresholds by Age



Note: This figure shows the effects of occupancy rate one day before patient's discharge on the likelihood of readmission by thresholds of length of stay (for the respective DRG and severity level) and by age, using the standard approach.

To shed some more light on what is behind the results, the regression in column 1 of Table 2.5 is repeated by aged groups, as defined previously in the model. Figure 2.4 shows the coefficients of occupancy rate one day before discharge and respective confidence intervals for each group of length of stay and each age group, in the standard approach. The respective graph for the conservative approach is in the appendix, in Figure B.4.

The positive coefficient for the maximum length of stay group is being driven by the younger patients, hinting towards some very extreme occurrences, such as accidents. The opposite happens on the above and below expected length of stay groups, since the older age groups are the ones driving the coefficients in their respective directions. The elderly are more sensible to length of stay, and discharging an older person before reaching the expected length of stay (conditional on DRG and severity) creates a risk in the form of readmissions. Insufficient attention to this group is driving more readmissions, causing inequalities in the access to inpatient care.

2.7 Other results

The large number of fixed effects is very important for identification purposes, but in the model they are absorbed for computational reasons. In this section I provide more information about time, hospital and MDC variables.

Appendix B.2, in the appendix, debates the effect that days of the week have on readmissions. The section shows that there is a clear weekend effect with occupancy and readmissions decreasing significantly in the weekends. Still, the effect does not translate into the relation between occupancy and readmissions. The effect of occupancy rates one day before discharge is positive for all days of the week. However, only the variables Tuesday, Wednesday and Friday present statistically significant results, which implies that occupancy on Mondays, Tuesdays and Thursdays are driving the results.

Table B.7 shows that occupancy rates one day before discharge are specially impactful in driving readmission episodes in the months of January, February, April and July, with positive statistically significant coefficients both in the standard and conservative approach. Occupancy variation one day before discharge affects readmissions mostly in January, March, April and June, when the coefficients are statistically significant in both the standard and the conservative approach.

So far, I also did not disentangle the analysis by MDC. It is likely that hospital occupancy has different effects on the readmission rates across health conditions. Table B.6, in Appendix B.2.1, shows that the positive effects of occupancy rates one day before discharge on readmissions are significant when considering conditions relating to the nervous, respiratory, circulatory, hepatic and urinary systems. The coefficients are larger in the respiratory and hepatic system. The digestive and musculoskeletal systems present positive but non-statistically significant coefficients.

Results may be further discriminated by hospitals characteristics. Appendix B.2.3 shows, in Figure B.5, Figure B.6 and Figure B.7, the differences between hospitals' region, types and size. The results in Section 2.5.2 are true across all hospital characteristics. The effects of occupancy rates one day before discharge on the probability of readmission are positive for all hospital regions, types and sizes. There are also no signs of heterogeneity between hospitals with different characteristics.

2.8 Conclusions

Portuguese Hospitals have very high occupancy rates. This fact has two implications. On one hand, it means that resources are not being wasted. On the other hand, they may not be sufficient to respond to patients' needs. I address the latter concern by performing a statistical quantification on the relation between occupancy rates and readmission episodes. Using data from the Portuguese DRGs for the years 2014, 2015 and 2016, I correlate occupancy variables experienced by patients during their hospitalization (in the form of occupancy rates and variation in occupancy rates) with future readmission episodes. I leverage on a high number of fixed-effects model to look at variation occurring within each week at each hospital.

The results show that occupancy rates, measured one and two days before discharge, lead to future readmission episodes. The effects are stronger when evaluated at higher rates of occupancy and for older individuals. Thus, it is fair to assume that the need to admit more people to a crowded hospital likely forces doctors to discharge other patients. The hypothesis of early discharges is confirmed using exogenous thresholds of length of stay, based on the Portuguese DRG pricing scheme, to conclude that occupancy variation and occupancy rates affect the ones that have length of stay below expected. Again, those are the ones belonging to an older age group, reinforcing the idea that early discharges predominantly affect a frailer segment of the population.

Part III

Informal long term care

Chapter 3

Partners in care! Health effects of providing care to spouses or partners

3.1 Introduction

Ageing of the population brings new challenges to developed countries. The increase in life expectancy comes with an increase in the number of years lived at latter stages of life and, consequently, frailty. The lack of capacity to perform everyday tasks, due to the decline in physical and mental capabilities, force the elderly to resort to the support of a third party: health professionals, family members, friends or neighbors. This, and the fact that most people age eventually, makes the phenomenon a societal issue that cannot be ignored. As such, policy makers around the globe have been trying to assess the best ways of providing care to the elderly.

A lot of attention has been given to informal care, specially since researchers confirmed it to be related to formal care, as a substitute or as a complement, depending on the situation (Van Houtven and Norton, 2004; Bolin, Lindgren and Lundborg, 2008; Bonsang, 2009). This revelation, and consequent necessity to compare both alternatives, motivated a plethora of studies on the pros and cons of informal care.

Informal caregiving can have positive altruism effects on caregivers. It is sometimes rewarding, as it conveys the feeling that the caregiver is needed, increasing their sense of pride and self-fulfillment (Pinquart and Sörensen, 2003; Brouwer et al., 2004, 2006).

Nonetheless, most of the findings dwell in the negative side of the activity. Taking care of a sick or fragile person forces a burden on the persons caring, who have to dedicate extra time of their lives to the act of providing help. Adding a new obligation to a daily routine of an individual may, in some cases, affect negatively her quality of live. Also, it is likely to weaken her health condition, if the routine becomes too overwhelming or if the health condition of the patient starts affecting the mental state of the carer (Zarit, Reever and Bach-Peterson, 1980; Schulz and Beach, 1999; Vedhara et al., 1999; Yaffe et al., 2002; Brouwer et al., 2004; Bobinac et al., 2010*b*; Do et al., 2015)). Not only that, but the simple fact that a person dear to the carer is sick is enough to create an emotional burden. In case of a family member, that effect is often referred in the literature as the family effect (Brouwer et al., 1999; Bobinac et al., 2010*b*).

The burden of caregiving manifests in diverse forms which stem from functional to cognitive impairments (Pinquart and Sörensen, 2003; Hiel et al., 2015). Physical problems have been found on men as a cause of continued caregiving (Coe, Harold and Houtven, 2009). In terms of psychological conditions, care recipients' behavior and mood disturbance can impact caregiver quality of life (Sewitch et al., 2004; Brouwer et al., 2004; Heger, 2017). For instance, a condition like depression is associated with poor caregiver quality of life (Sewitch et al., 2004).

This paper is in line with studies evaluating the hazards of informal care. I try to assess the consequences of providing personal care on the health outcomes of a very particular type of carers: partners or spouses. Partner caregivers are a very special group in their characteristics. Unlike other informal providers, spouses and partners tend to be about the same age as the care receiver and live under the same roof. The type of tasks they perform exclusively as care providers, as well as the associated burden, may differ from conventional children to parents informal caregiving (Llácer et al., 2002). The care provided by spouses is an important mechanism which helps avoiding institutionalization, substituting or delaying formal long-term care (Nihtilä and Martikainen, 2008; Bakx and de Meijer, 2013). But co-residential care can lead to worst quality of life and higher number of depressive symptoms (Barbosa and Matos, 2014).

I use data from the waves 1, 2, 4, 5 and 6 of the Survey of Health, Ageing and Retirement in Europe (SHARE), for people people aged 50 or older, to select a subsample of potential partner caregivers based on the questions: "*Is there someone living in this household whom you have helped regularly during the last twelve months with personal care, such as washing, getting out of bed, or dressing?*" and "*Who is that?*". The dataset presents a matching code for partners, therefore allowing me the use partners' variables for each observation. I look at single living couples and compare partner caregivers with non-caregivers. The comparison is made using two distinct health outcomes. I consider physical health and mental health. Physical health is measured by a self reported measure and by a more robust disability index inspired by Bonsang 2009. For depression, I use the EURO-D[1] scale reported in SHARE.

I cannot perform the analysis without considering the possibility of family effects (Brouwer et al., 1999; Bobinac et al., 2010*a*). Moreover, since I am looking at partners, there are other likely health correlations across caregivers and care receivers. For instance, it is reasonable to think that the Grossman Model (Grossman, 1972) can be extended to couples who live together. Partners may share health investment decisions and health behaviors. Thus, people with lower health may be married to people with lower health, frailer individuals may be married to frailer individuals, etc... Therefore, partner caregivers are likely to be a selection of people with lower health, while still enough to be caregivers. One typical solution to this problem is to use an arbitrary occurrence that forces the individual in or out of caregiving. But sometimes those exogenous shocks with no correlation with the characteristics of the partner caregivers are difficult to find, limited in the number of observations, or just not available in the data. For that end, I include the health

[1]It is a depression scale present in the SHARE survey which is based on 12 objective questions about mental health.

characteristics of the patient in the model, in a first step, to help mitigating the health correlation selection problems. Then, I perform an event study with pre-treatment matching (Schmitz and Westphal, 2015) to study quasi-random transitions into spousal caregiving.

The static approach consists of regression analysis and matching. Due to the categorical nature of the dependent variables, the first approach uses an ordered probit model for self assessed health (SAH) and depression, and a generalized linear model with a probit link for the disability index. The estimated coefficients on the partner providing personal care are all indicative of worst self reported physical health and depression scales, when partner characteristics are not taken into account. However, the inclusion of those variables, causes relevant changes in the results. In the more complete methodology, partner caregivers display positive results on all health outcomes, when compared with similar non-caregivers who have equally sick partners. The results show no evident pattern across countries. They are robust to a matching strategy. Still, this approach only proves that caregivers are better off, but it does not imply causality.

To that end, the event study compares individuals who transition into caregiving with the ones who remain non-caregivers across waves of the survey. I match observations of both groups in the pre-treatment period, on all individual and partner characteristics used across this study, to improve the likelihood that transitions are random. The results of the subsequent regressions show only positive results for self perceived physical health, and non-conclusive results for mental health and disability index. Also, continued spousal caregiving has no impacts on providers health.

My analysis enters the literature of caregiver burden, particularly in the scope of low intensity co-residential care. Its contribution comes from showing that household investment decisions on health must be taken into account when searching for potential burden of partner caregiving. It provides information to policy makers, saying that partner caregiving has positive effects on self perceived physical health, and no negative effects on mental health or disability. So, they should aim at understanding what drives the positive results, e.g. sentiment of self-fulfillment or changes in reporting patterns, and promote partner caregiving under the right setup.

In Section 3.2 the paper presents the data, in Section 3.3 the methodology, in Section 3.4 the results and in Section 3.5 it looks at changes across waves. Section 3.6 concludes.

3.2 Data and summary statistics

I use data from the Survey of Health, Ageing and Retirement in Europe (SHARE), a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks for individuals aged 50 or older. SHARE covers 27 European Countries through 7 available waves of questionnaires[2], at the time of this study. I use waves 1, 2, 4, 5 and 6, collected in 2004, 2007, 2011, 2013 and 2015, respectively.

To understand whether partner caregivers have different health outcomes due to their caring activities, I define two groups, individuals who provide personal care to their partners (treatment) and individuals who do not provide any type of help to their partners (control). Single person households are excluded. Partner caregivers are determined by the question: "*Is there someone living in this household whom you have helped regularly during the last twelve months with personal care, such as washing, getting out of bed, or dressing?*"[3] and "*Who is that?*". If the individual receiving help is the partner or the spouse, then I classify the individual as a partner caregiver. I identify all other care provided at home, by checking if the partner received any formal (paid) care or informal care from people outside the household or other household members and if the care provider helped anyone else other than the partner. The survey provides a partner identification code, that allows me to match partners. Also, partners are necessarily interviewed, guaranteeing I have access to their answers.

Regarding the outcome variables, I look at physical and mental health outcomes, using physical Self Assessed Health (SAH), Depression Scale and Disability Index. The SAH is based on the

[2]Not all countries are present in all waves.

[3]The variable is binary (yes or no) and there is no information about the amount of time spent providing personal care.

question "how do you rate your physical health on a scale from 1 to 5". The Depression Scale (Prince et al., 1999) ranges from 1 to 12 and is based on the evaluation of the following states of mind: depression, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness. As a robustness check for physical health, I built a disability index by running an Ordered Probit regression of the level of limitation (self reported) on characteristics of disability (instrumental activities of daily living) and standardizing the predicted values of the latent variable. This method alleviates self report bias that is intrinsic to variables such as self assessed health (Bonsang, 2009). Thus, it works as a sanity test.

The sample has 99,613 observations. The summary statistics for the main variables of the study are displayed in Table 3.1. They compare individual statistics and partner statistics for both control and treatment group.

Table 3.1: Summary Statistics

	Caregivers	Non-Caregivers	Min	Max
Self assessed health	2.49	2.94	1 (very bad)	5 (very good)
Depression scale	8.87	9.85	0 (very depressed)	12 (not depressed)
Disability index	0.75	0.83	0.07 (very limited)	0.98 (not limited)
Years of schooling	10.35	10.97	0	25
Age	69.99	66.20	24	99
Gender (male)	46.03%	50.20%	0%	100%
Partner caregiver	100%	0%	0%	100%
Employed	10.65%	23.41%	0%	100%
Retired	71.73%	61.15%	0%	100%
Homemaker	11.05%	9.50%	0%	100%
Sick or disabled	3.77%	2.75%	0%	100%
Make ends meet	2.69	2.99	1 (with great difficulty)	4 (easily)
Partner's characteristics:				
Self assessed health	2.00	2.97	1 (very bad)	5 (very good)
Depression scale	8.87	9.85	0 (very depressed)	12 (not depressed)
Disability index	0.75	0.83	0.07 (very limited)	0.98 (not limited)
Years of schooling	10.24	10.97	0	25
Age	70.84	66.16	24	99
Observations	5,221	94,392		

The table displays the summary statistics separated in two groups, partners caregivers and non-caregivers. The first two columns show the average values of each variable, while the two last columns show the maximum and minimum values, respectively.

The table shows differences between partner caregivers and non-caregiver. Caregivers have worse indicators in all variables. They are older, have lower levels of physical and mental health, lower education, higher likelihood of being sick or disabled and more difficulty in making ends meet. They are also the ones with sicker partners, as shown in the partner's characteristics section of the table.

The difference in some indicators may be explained by selection rather than caregiver status because the need for personal care is more likely to occur for people of more precarious social and economic backgrounds. Those are the ones who might have "opted" for lower investments in health in their lives.

The summary statistics for the entire sample are provided in annex, in Table C.1.

3.3 Methodology

3.3.1 The baseline model

I use the following basic specification:

$$\text{Health}_{ijt} = \beta.\text{Care}_{ijt} + \gamma.X_{ijt} + \theta.\text{Country}_j + \eta\text{Wave}_t + \varepsilon_{ijt} \tag{3.1}$$

The outcome variable, Health_{ijt}, is one of three: a measure of self reported physical health, a measure of disability or a measure of mental health, for individual i of country j in wave t. The physical outcome, given by the self assessed health, is an ordinal variable with discrete values ranging from 1 to 5, where the highest number corresponds to a very good physical health. The disability measure is an index, between 0 and 1, with 1 corresponding to a perfect state with no disabilities. The depression scale is an ordered variable whose discrete values go from 1 to 12, with 12 corresponding to a perfect mental healthy state.

The main explanatory variable, Care_{ijt} is a binary variable taking value 1 if the individual provides personal care to the corresponding partner and 0 otherwise.

Vector X_{ijt} includes individual characteristic such as gender, age, education, ability to make ends meet and employment status. It also includes binary variables stating whether any other type of care is provided in the household of the potential caregiver[4]. Depending on the specification, it may also include partner characteristics, such as self assessed health, disability index, depression scale, education and age. Finally, the model accounts for Country_j and Wave_t fixed effects.

Since both self assessed health and depression scale are ordinal, I perform ordered probit estimations which I complement with the simpler ordinary least squares estimations. For the disability index, I compute a generalized least squares with a probit link.

In this paper, I compare two alternative specifications of the previous model. First, I estimate the model with only individual and household characteristics. Then, I redo the regressions to account for the partners' health condition. The difference between the coefficients of caregiving for a partner are attributed to household effects on health, more specifically, households' health investment decisions. The idea behind this methodology is that the health production function depends not only on the individual investment in health but rather on a joint decision with the other household members. Considering investments in health as direct investments in health care and prevention or adoption of risky behaviors means, that with higher probability, a frailer individual belongs to a household where the investment in health was chosen to be lower. So, the health of the individual should be affected by the health of the partner, since they took health decisions together. Additionally, sickness and frailty of the patient can have direct effects on the health of close relatives through the family effects. Therefore, the partner of a sick individual may suffer emotionally and consequently report lower health outcomes.

[4]Includes dummies for providing care to other individual who are not partners and receiving help from other individuals who are not partners

3.3.2 Partner caregiving by country

Another important part of my analysis is to discriminate how partner caregivers fare in different countries. The differences in the approaches to long term care and the cultural traits of each country have the potential to shape the sample in analysis. For example, if there are countries where long term care absorbs the most severe cases, only the lower intensity cases will be left for partners to treat. Or alternatively, if a country provides more support for partner caregivers, it may attenuate the negative effects and foster the positive effects. That type of selection is not discriminated in the overall results, which present an average effect. But I am questioning whether coefficients across countries are the same, and if so, which ones drive the results. To that end, I repeat the analysis in Section 3.3.1, but separating the personal caregiving effect for each country.

3.3.3 Propensity score matching

Given that the sample is unbalanced, with big differences between treatment and control group I apply a matching estimator to guarantee a more comparable control group. Implementing a one to one propensity score matching estimator allows me to compare treated with non-treated individuals, that are similar on a set of observed characteristics.

The matching is done conditioning on the following set of covariates: partner's self assessed health, partner's depression scale, partner's disability index, partner's education and partner's age. For the individual characteristics, I match on individual age, gender, education, employment, household size, ability to make ends meet, country and wave.

The propensity score matching estimator allows for a direct comparison between treatment and control group of similar observed characteristics in a balanced fashion, alleviating the selection problems of the sample. In this case, matching answers two questions: 1) What would be the health outcomes of non-caregivers if their partners were as sick as the partners of the caregivers? 2) What is the difference in health outcomes between non-caregivers and partner caregivers if their partners have similar health conditions?

Note that matching does not solve the potential problem of reverse causality. Individuals in the matched group have partners with similar health conditions, thus, they could be providing them personal care. If the factor explaining that decision is not altruism, then it is possible that they selected themselves into not providing personal care because they don't feel they are physically or mentally able to perform those tasks. Therefore, their health outcomes would be lower due to selection.

3.3.4 Event Study

Performing an event study on the health effects of becoming a partner caregiver eliminates the risk of finding results tainted by reverse causality, if transition into caregiving is random. I compare the variation in health of individuals who transition into spousal caregiving with individuals who remain non-caregivers. Still, it is likely that those two groups do not have the same rates of health depreciation, due to their differences in characteristics. To account for that, I match them on observables before treatment, i.e., when individual on both groups are still non-caregivers. If both groups are non-caregivers, have the same health, education, income, age, gender and equally healthy/sick partners, in the first wave they are observed, then transitions into caregiving are likely to occur due to deterioration of their partners' health. The transition into caregiving is therefore more likely to be random, instead of a product of group selection (Schmitz and Westphal, 2015).

3.4 Results Static approach

3.4.1 Physical health

In Table 3.2, I show the results for self assessed health. In the first two columns, without the inclusion of partners' variables, the coefficients of both the OLS and the OProbit are negative and significant, indicating that partner caregivers have a lower health status than non-caregivers. However, in the third and fourth column, when partners' variables are added to the model, the coefficient's direction switches. Being statistically significant, it says that when caregivers and non-caregivers have similar spouses or partners, the caregivers show better self reported physical health. This implies that household selection components are present in the model, and the inclusion of partner characteristics helps mitigating that type of selection[5].

The average marginal effects for the ordered probit model concerning self reported health are shown in Figure 3.1. When no partner variables are taken into account (OProbit 1) the individual who provides personal care to a partner is 1 percentage point more likely to be in outcome 1, and 1.5 percentage points more likely to be in outcome 2, while roughly 1.5 percentage points less likely to be in outcomes 4 or 5. When the partner's variables are used in the estimation procedure (OProbit 2), the figure rotates slightly before outcome 3. Now, the carer is 1 and 1.5 percentage points less likely to be in outcomes 1 and 2, respectively. Outcome 4 and 5 are now about 1 percentage points more likely.

[5] But it may create another type of selection driven by reverse causality. This issue is only addressed in the event study section.

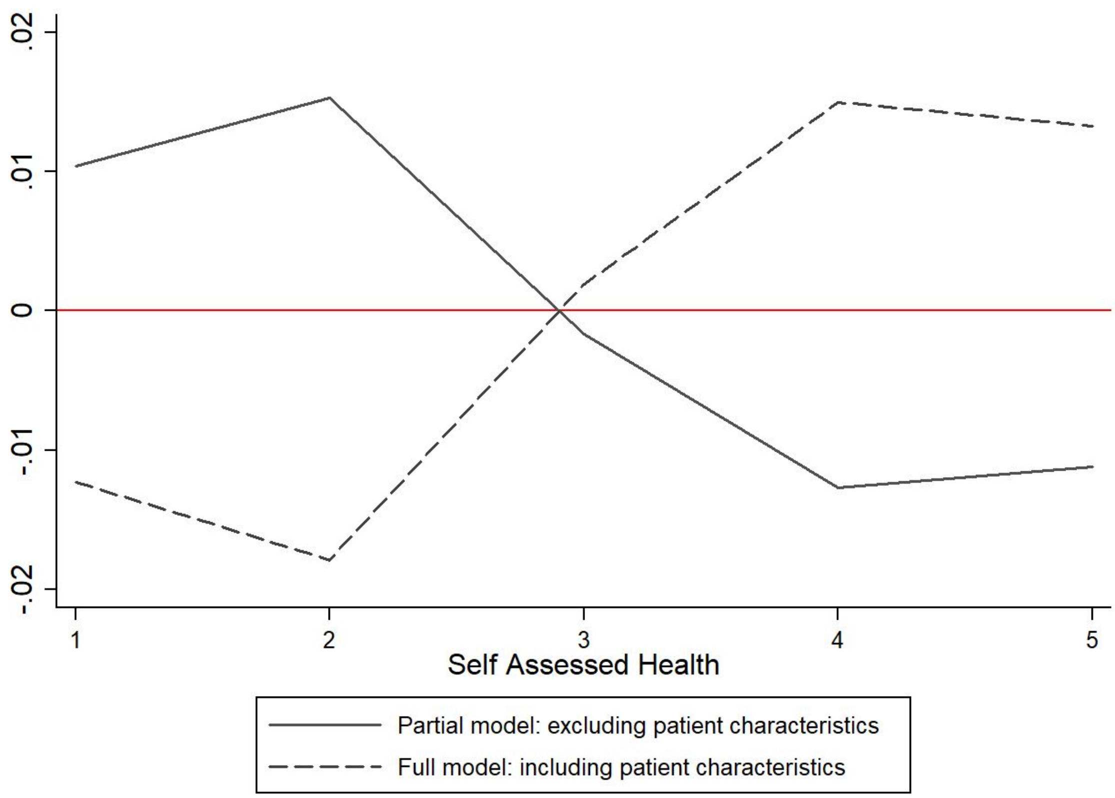
Table 3.2: Main estimates on Self Assessed Health

Dependent variable: Self Assessed Health	OLS	OProbit	OLS	OProbit
	w/o partners' characteristics		w/ partners' characteristics	
Partner caregiver	-0.075***	-0.084***	0.075***	0.102***
	(0.014)	(0.017)	(0.015)	(0.018)
Partner's self assessed health			0.171***	0.201***
			(0.006)	(0.007)
Partner's depression scale			0.019***	0.024***
			(0.002)	(0.002)
Partner's disability index			-0.015	0.012
			(0.031)	(0.039)
Partner's age			0.026***	0.033***
			(0.005)	(0.007)
Gender	0.045***	0.051***	0.062***	0.075***
	(0.007)	(0.008)	(0.009)	(0.011)
Age	0.010*	0.015*	-0.008	-0.007
	(0.005)	(0.006)	(0.006)	(0.007)
Make ends meet	0.161***	0.192***	0.120***	0.146***
	(0.004)	(0.005)	(0.004)	(0.005)
Employed or Self-Employed	0.192***	0.219***	0.192***	0.224***
	(0.011)	(0.013)	(0.011)	(0.013)
Unemployed	-0.088***	-0.099***	-0.069**	-0.078**
	(0.024)	(0.028)	(0.023)	(0.027)
Permanently sick or disabled	-0.808***	-1.034***	-0.769***	-1.008***
	(0.020)	(0.028)	(0.020)	(0.028)
Homemaker	-0.002	-0.005	-0.021	-0.028
	(0.013)	(0.016)	(0.013)	(0.016)
Other Status	0.017	0.015	0.018	0.017
	(0.032)	(0.038)	(0.031)	(0.037)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regressions use a total of 99,613 individual surveys collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 20 countries. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. “Retired” is the base for the employment factor variables. Other type of care (provided or received), age squared and degree of education of individuals and partners are also included as control variables, but not displayed in the table.

Figure 3.1: Partners in Care: Average Marginal Effects on SAH



The figure shows the average marginal effects of providing care to a partner in need on the self assessed health of the carer. The partial model takes no partner variables into account. In this model individuals who provide personal care to a partner are more likely to report lower physical health outcomes and less likely to report higher physical health outcomes. The full model has partners' variables into account. In this model individuals who provide personal care to a partner are less likely to report lower physical health outcomes and more likely to report higher physical health outcomes.

3.4.2 Depression

In Table 3.3, I show the results for depression. Columns 1 and 2 show the OLS and OProbit coefficients for providing care to a partner when no partner characteristics are introduced. The values are negative and statistically significant, suggesting that partner caregivers have worse mental health. Again, when partners' characteristics are introduced, columns 3 and 4, the coefficients change. In this case, they lose in magnitude and become positive, but are not statistically significant. Thus, we cannot reject the hypothesis that partner caregivers and non-caregivers have similar depression levels.

The average marginal effects of caregiving when depression is the outcome are illustrated in Figure 3.2. In this figure, when ignoring the partner variables, all the outcomes bellow ten are more likely to occur if the individual is a partner caregiver, with the highest average marginal effects being around 1.5 percentage points for outcome 8. The partner caregiver is less likely to be in the two better outcomes, with 12 having a decrease in probability of almost 6 percentage points. When including partner variables, the figure is very close to a straight line, with only a small increase in probability for the two last outcomes.

Regarding the other coefficients, both Table 3.2 and Table 3.3 display similar results in terms of direction. The partners' health coefficients for physical self reported health, depression scale and disability index are all positive, demonstrating that there is a significant positive correlation between partners' health outcomes. Concerning the other variables, it is noteworthy that men are healthier than women and that ability to make ends meet is associated with better health[6] (in both outcomes). Employment status possibilities also present the expected signs. Only age presents a variation across tables since it is positively associated with depression scale, while it is negatively associated with physical health.

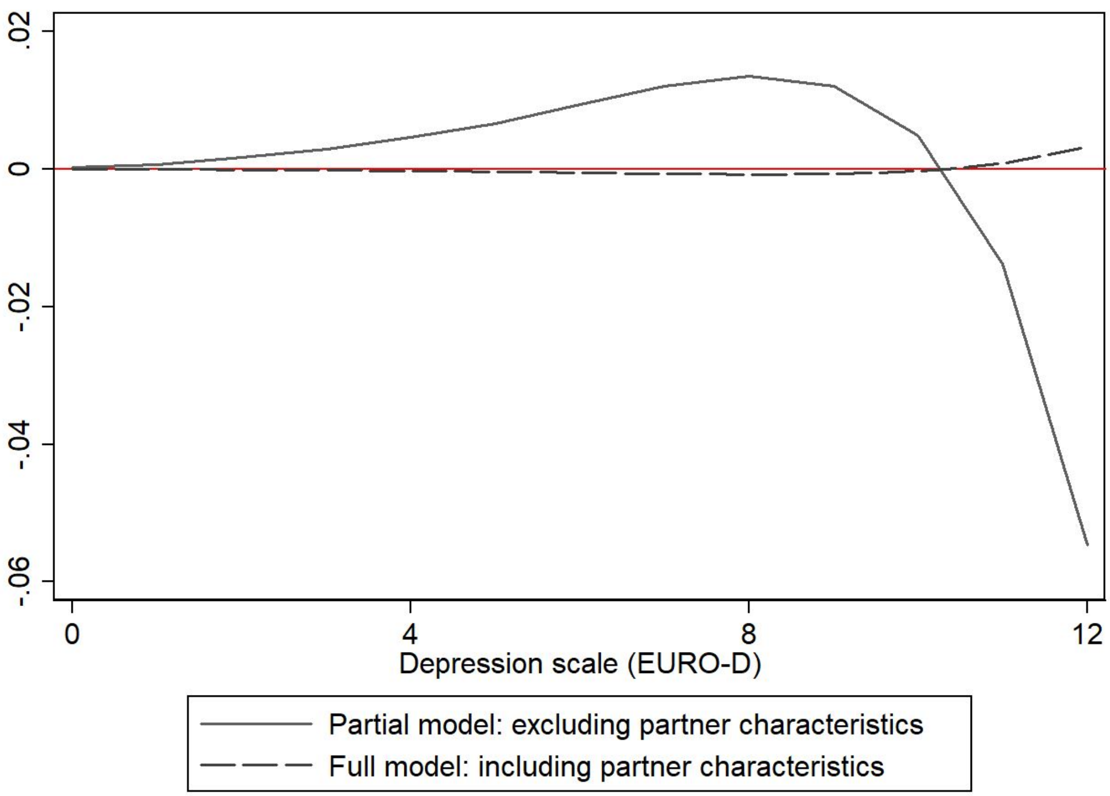
[6] In line with the health production models

Table 3.3: Main estimates on Depression Scale

Dependent variable: Depression Scale	OLS	OProbit	OLS	OProbit
	w/o partners' characteristics		w/ partners' characteristics	
Partner caregiver	-0.369***	-0.188***	0.012	0.011
	(0.034)	(0.016)	(0.035)	(0.017)
Partner's self assessed health			0.015	0.013**
			(0.008)	(0.005)
Partner's depression scale			0.265***	0.144***
			(0.005)	(0.003)
Partner's disability index			0.066	0.010
			(0.072)	(0.037)
Partner's age			0.019	0.004
			(0.012)	(0.006)
Gender	0.715***	0.390***	0.932***	0.526***
	(0.014)	(0.008)	(0.020)	(0.011)
Age	0.118***	0.061***	0.083***	0.048***
	(0.011)	(0.006)	(0.012)	(0.006)
Education degree	0.035***	0.019***		
	(0.003)	(0.002)		
Make ends meet	0.343***	0.173***	0.233***	0.120***
	(0.010)	(0.005)	(0.008)	(0.004)
Employed or Self-Employed	0.079***	0.049***	0.102***	0.061***
	(0.023)	(0.013)	(0.021)	(0.013)
Unemployed	-0.359***	-0.195***	-0.323***	-0.186***
	(0.054)	(0.026)	(0.051)	(0.026)
Permanently sick or disabled	-1.174***	-0.561***	-1.081***	-0.536***
	(0.053)	(0.024)	(0.050)	(0.023)
Homemaker	-0.124***	-0.051***	-0.155***	-0.071***
	(0.031)	(0.015)	(0.030)	(0.016)
Other Status	-0.199**	-0.103**	-0.171**	-0.095**
	(0.069)	(0.035)	(0.064)	(0.034)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Regressions use a total of 99,613 individual surveys collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 20 countries. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. "Retired" is the base for the employment factor variables. Other type of care (provided or received), age squared and degree of education of individuals and partners are also included as control variables, but not displayed in the table.

Figure 3.2: Partners in Care: Average Marginal Effects on Depression



The figure shows the average marginal effects of providing care to a partner in need on the depression scale of the carer. The partial model takes no partner variables into account. In this model individuals who provide personal care to a partner are more likely to report lower depression outcomes and less likely to report highest depression outcomes. The full model accounts for partners' variables. In this model individuals who provide personal care to a partner are as likely to report lower depression outcomes as individuals who do not provide care, but slightly more likely to report the highest values of the depression scale.

3.4.3 Disability

When using self reported measures, the bias inherent to the answers can be a major problem in econometric analysis. Factors such as education, income and others may drive different groups to answer in different ways (Bago d'Uva, O'Donnell and Van Doorslaer, 2008). I may find that problem when using physical self-reported health.[7] One way to attenuate the bias for such variables is to use a list of predictors to estimate its latent variable and standardize the predicted values to create an index. One alternative commonly used to substitute the self assessed health is the disability index. I built the disability index using a categorical question on how much the individual feels limited in activities because of her health. I regressed limitation on a group of binary variables about activities of daily living and health problems using a probit model. With the estimated coefficients I determine the maximum and the minimum predicted value for disability, necessary to standardize the individual predicted values. This disability index is therefore an "anchored" measure, less prone to bias and measurement error.

Table 3.4 shows the ordinary least square and the generalized linear model (with a probit link) estimations for disability. The numbers in the GLM columns correspond to the average marginal effects. This table tells the same tale as Table 3.2. When no partner caracteristics are included in the model, caregiving has a negative sign. However, it turns positive when partner's health, education and age are taken into account. In the fourth column, the average marginal effects tell that an individual who is a partner caregiver has, on average, two extra percentage points in the disability index. Therefore, partner caregivers feel slightly less limited in activities.

Note that there is a strong positive correlation between the level of disability of the individual and the one of the correspondent partner. Again, it supports the theory that health investment decisions are made at the household level.

[7] The same may be true for reports of depression levels, but on a lower scale since the Euro-D is calculated using the answers to twelve objective questions.

Table 3.4: Main estimates on Disability Index

Dependent variable: Disability Index	OLS	GLM	OLS	GLM
	w/o partners' characteristics		w/ partners' characteristics	
Partner caregiver	-0.011***	-0.008***	0.019***	0.020***
	(0.002)	(0.002)	(0.002)	(0.002)
Partner's self assessed health			0.000	0.001*
			(0.000)	(0.001)
Partner's depression scale			0.003***	0.002***
			(0.000)	(0.000)
Partner's disability index			0.152***	0.124***
			(0.006)	(0.006)
Partner's age			0.004***	0.003***
			(0.001)	(0.001)
Gender	0.032***	0.032***	0.039***	0.039***
	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.007***	0.005***	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Make ends meet	0.020***	0.020***	0.015***	0.015***
	(0.001)	(0.001)	(0.000)	(0.001)
Employed or Self-Employed	0.010***	0.018***	0.012***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)
Unemployed	-0.004	-0.001	-0.002	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)
Permanently sick or disabled	-0.117***	-0.109***	-0.113***	-0.105***
	(0.003)	(0.004)	(0.003)	(0.003)
Homemaker	-0.003	-0.002	-0.005**	-0.004*
	(0.002)	(0.002)	(0.002)	(0.002)
Other Status	-0.004	-0.002	-0.003	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)

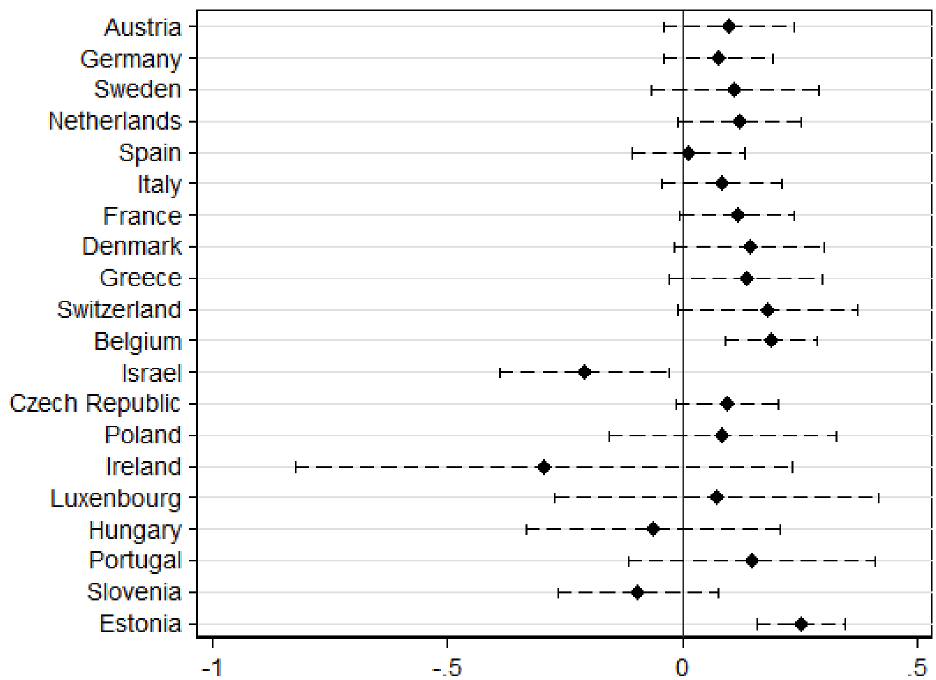
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table displays the average marginal effects. Regressions use a total of 99,613 individual surveys collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 20 countries. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. "Retired" is the base for the employment factor variables. Other type of care (provided or received), age squared and degree of education of individuals and partners are also included as control variables, but not displayed in the table.

3.4.4 Partner caregiving effect by country

The results discriminated by country are presented in Figure 3.3, Figure 3.4 and Figure 3.5. The figures show the interaction of providing personal care to a partner and country, following the methodology of Equation (3.1).

Figure 3.3: SAH of partner caregivers

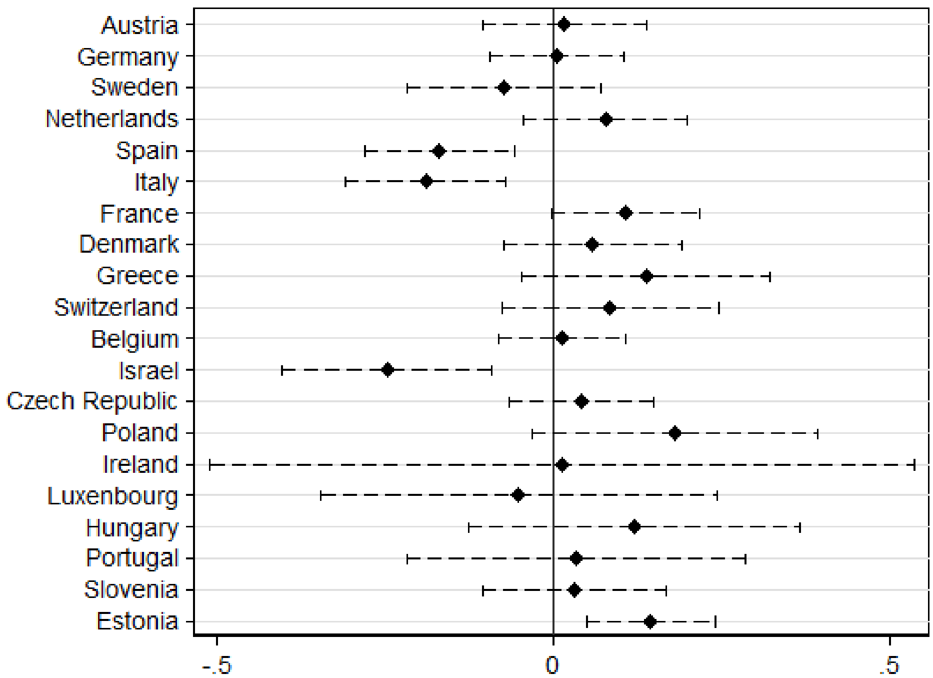


The graph shows the coefficients of the variable "Partner caregiver" discriminated by country, in the self assessed health regression. The model includes both individuals and partner characteristics.

Figure 3.3 shows the discriminated results for self assessed health. Providing personal care to a partner in need is mostly positive across countries, however, only statistically significant for Belgium and Estonia. Only Slovenia, Hungary, Ireland and Israel present negative point estimates. Only Israel shows significance. Note that heterogeneity is weak across the board. If Belgium and Slovenia are left out, no country is statistically distinguishable from any other.

Figure 3.4 displays the coefficients of partner caregiving on the depression of the carers by country.

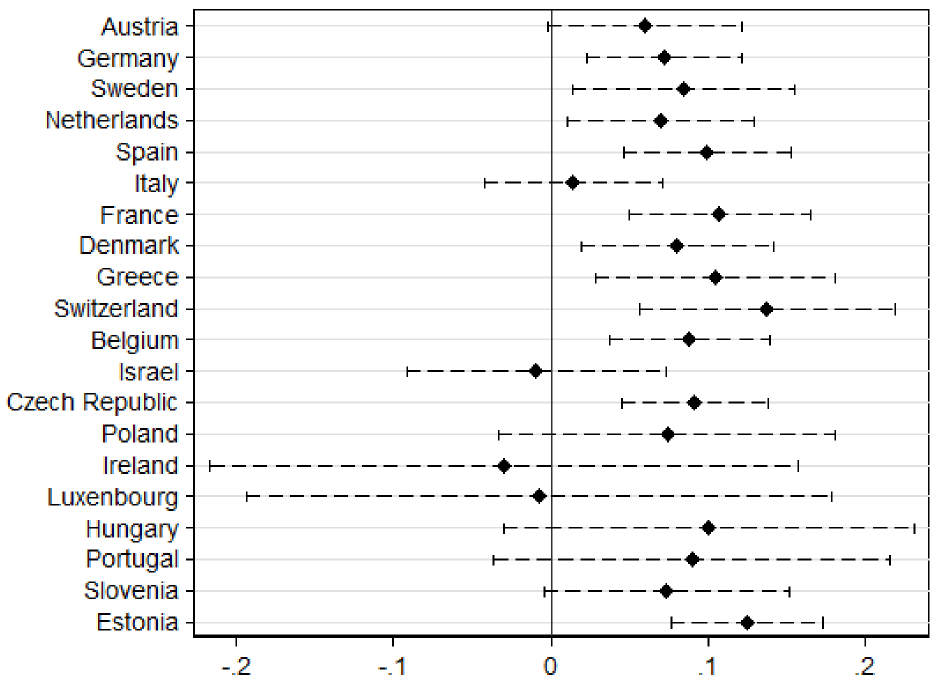
Figure 3.4: Depression of partner caregivers



The graph shows the coefficients of the variable “Partner caregiver” discriminated by country, in the depression regression. The model includes both individuals and partner characteristics.

Again, most countries show positive point estimates (15), but only one with statistical significance. The remaining five countries have negative point estimates, with Spain, Italy and Israel presenting coefficients statistically distinguishable from zero. Similar to the SAH analysis, heterogeneity is also weak for depression. Besides Spain, Italy and Israel, countries are not statistically distinguishable between each other.

Figure 3.5: Disability of partner caregivers



The graph shows the coefficients of the variable "Partner caregiver" discriminated by country, in the disability regression. The model includes both individuals and partner characteristics.

Figure 3.5 displays the country coefficients of spousal caregiving for the disability outcome. Only three countries (Ireland, Luxembourg and Israel) present negative point estimates, with no statistically significance. All the remaining countries, have positive point estimates for partner care. From those, eleven present statistical significance.

This section shows that when accounting for partner characteristics, partner caregiver present better levels of physical health (mainly though their levels of disability) in most of the countries in analysis. It also shows that results from depression are more mixed. In the latter, caregivers from Spain, Italy and Israel present clear losses when compared to non-caregivers. There appear to be no major patterns in terms of regions. There is no heterogeneity between countries since no country can be statistically distinguishable from any other.

3.4.5 Other results - heterogeneity

The literature on informal care has shown differences across gender, namely on the likelihood of becoming informal caregivers. In this section I want to define whether there is a difference in health between male and female partner caregivers. Figure C.1, Figure C.2 and Figure C.3 show the results of the coefficients of being partners caregivers on SAH, depression and disability, discriminated by gender. Regarding the two physical measures, SAH and disability index, both male and female caregivers are more likely to have better outcomes. However, there is no heterogeneity since they are not statistically distinguishable. The depression coefficients present a positive point estimate for men and negative for women, but again, not statistically distinguishable from each other.

Results may not be linear in health. The difference between partner caregivers and non-caregivers may depend on the starting level of health. Therefore, I also look at the coefficients of being a partner caregiver on health for three different levels of disability. I divide the analysis into three groups: individuals with disability index higher than 90%; in between 80% and 90%; and below 80%. The results for SAH, depression and disability are presented in Figure C.4, Figure C.5 and Figure C.6, respectively. The results for the two physical health measures show higher partner caregiver coefficients for the more physically able individuals. This indicates that the difference between caregivers and non-caregivers is higher for those with better physical capabilities. Note, that in Figure C.6, the two better physical health groups are statistically distinguishable from the worst group, showing some heterogeneity. In terms of depression, in neither groups do caregivers present differences from non-caregivers. There is also no heterogeneity between groups.

3.4.6 Matching

3.4.6.1 Determinants of partner care

The results of the Probit model implicit to the propensity score matching are shown in Table 3.5. As expected, the probability of providing care to a partner depends negatively on all three health outcomes of the partner. No one is going to offer (demand) personal care if personal care is not needed. Thus, lower health outcomes of the potential patient are a necessary condition for caregiving to happen. The age of the individual negatively affects the probability of providing personal care while the age of the partner naturally increases it. Additionally, women are more likely to be partner caregivers, and higher capacity to make ends meet is associated with higher probability of caregiving. Employment status says that employed individuals are less likely to provide care to a partner when compared to retired individuals.

I choose the variables in a way to minimize unobserved heterogeneity. The combination of the individual's variables with the partner variables allows me to compare two groups that face the same hardships of frailty and disease of the partner.

Health outcomes of the potential carer cannot be used in the matching process, as they are the variables of interest that are going to be compared across groups. But since health of the potential carer may also be a predictor of informal caregiving there might be room for reverse causality. I address that issue in the event study section.

In Table C.2, in annex, I display the means for the treated and matched group obtained through the propensity score matching. Note that this is a one to one matching with a caliper of 0.2 of the standard deviation. The values in the table are very close for all variables, guaranteeing that the two groups are comparable. For instance, the partner disability index is, on average, 83% for the treated, against 83.2% for the matched, which implies that patients and potential patients have similar capacities for performing daily activities. Alternatively, there may be reverse causality coming from an unobserved “caring ability”? Maybe caregivers selected themselves into caregiving be-

Table 3.5: Propensity Score Matching - Determinants of caregiving

Dependent variable: Partner Caregiver	Probit
Partner's self assessed health	-0.131***
	(0.011)
Partner's depression scale	-0.017***
	(0.004)
Partner's disability index	-2.949***
	(0.067)
Partner's age	-0.042**
	(0.014)
Gender	-0.148***
	(0.021)
Age	0.049***
	(0.014)
Make ends meet	0.003
	(0.009)
Employed or Self-Employed	-0.119***
	(0.031)
Unemployed	-0.033
	(0.059)
Permanently sick or disabled	-0.064
	(0.049)
Homemaker	-0.010
	(0.030)
Other Status	0.098
	(0.077)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table displays the Probit model implicit to the computation of the propensity scores. It uses a total of 99,613 observations. It includes fixed effects at the country and wave level. "Retired" is the base for the employment factor variables.

cause they have a natural aptitude for giving care, or because they have better relations with their partners? If such an unobserved characteristic is related with health then it would cause biased estimates.

3.4.6.2 Average treatment effect on the treated

In Table 3.6, I show the averages of the treated, matched control and unmatched control group. The "unmached row", displays the difference in health outcomes before matching. Both depression, self assessed health and disability are reported at lower values for partner caregivers. This is similar to the results found in the "no partners' variables" regressions. When the groups are comparable however, the treated have slightly better health outcomes than the matched control group. This is an important result because it states that the caregivers are not worse than non-caregivers, when their partners condition is of similar severity. Nevertheless, it seems that what causes worse outcomes is very likely to be their household decisions of health investment and consequent correlation of health. This correlation may be due to risky habits and other behavioral components. Also, the emotional burden of having a sick or frail partner may impact the health of the individual.

One concern when interpreting these results is the potential existence of reverse causality. Indeed, one reason that might explain the difference in caregiver status between individuals with similar observed characteristics may be that their health condition does not allow them to be caregivers. In that case, the regression and matching results would be capturing this effect. Then, caregivers would not be healthier because they are caregivers, but instead, they would be caregivers because they are healthier (in relation to the matched control group).

Table 3.6: Matching Estimates: Effects of Providing Care to a Partner

Self Assessed Health	Treated	Control	Difference	Standard Error
Unmached	2.489	2.947	-0.457	0.015
Matched	2.504	2.433	0.072	0.021

Depression Scale	Treated	Control	Difference	Standard Error
Unmatched	8.874	9.864	-0.989	0.030
Matched	8.922	8.842	0.080	0.050

Disability Index	Treated	Control	Difference	Standard Error
Unmatched	0.752	0.830	-0.077	0.002
Matched	0.756	0.723	0.033	0.003

The matching uses a sample of 97,89 individuals, comparing a treatment group of 5,082 persons with a control group of 92,811 persons. The results illustrate a 1 to 1 propensity score matching with no replacement and a caliper of 0.2 of the standard error, in the selection equation. The analysis is performed on the common support.

3.5 Event study - Changes in caregiving status

3.5.1 The model

So far I looked at a static problem, to conclude that, at a given point in time, caregivers have better reported health measures, when the analysis is conditioned on partners' health characteristics. In this section, I address the changes in caregiver status across waves. I perform an event study with all waves in analysis to capture transitions of non-caregivers into caregivers. The subsample includes two groups: non-caregivers who become caregivers in the subsequent wave and non-caregivers who remain non-caregivers. I apply the same line of thought I used throughout this paper to a dynamic framework. I include the partners' characteristics in a differences in differences model which is written in the following way:

$$\begin{aligned}\text{Health}_{ijt} = \beta_0 + \beta_1\text{Period2}_t + \beta_2(\text{NC} \rightarrow \text{C})_i + \beta_3[\text{Period2}_t * (\text{NC} \rightarrow \text{C})_i] \\ + \gamma.X_i + \theta.\text{Country}_j + \varepsilon_{ijt}\end{aligned} \tag{3.2}$$

$Health_{ijt}$ is one of the three health outcomes used in this paper. $Period2_t$ is a binary variable taking value 1 if the observation belongs to the second wave in which the individual is observed (i.e for an individual who transitions into caregiving it corresponds to the period where she is a caregiver). The variable $(\text{NC} \rightarrow \text{C})_i$ is a binary variable which takes value 1 if the individual transitions into caregiving.

X_i is a vector which includes individual and partner characteristics and Country_j stands for country fixed effects. ε_{ijt} is the error term.

Note, however, that studying transitions into spousal caregiving is not enough to guarantee causality. Individuals of the two groups may have different health depreciation rates, independently of the change in caregiver status. For example, caregivers may be part of households with lower health investments (e.g. behavioral risks), and thus, present higher depreciation of health. Alternatively, if reverse causality exists, after controlling for partner characteristics, then non-caregivers are likely to be the ones with lower depreciation rates. To account for this, I perform a propensity score matching between the groups in the first period they are observed (i.e, individuals of both groups are still non-caregivers). The regression is then performed on the common support restricted sample. This regression adjusted matching approach is inspired by Schmitz and Westphal 2015. By forcing individuals to be "equal" in the pre-treatment period, I reduce the probability that health variations are driven by different depreciation rates across groups. Now, reverse causality is unlikely since the health of both non-caregivers and caregivers is identical pre-treatment. In this framework, transitions into spousal caregiving are more likely to be random.

3.5.2 Results

Table 3.7 shows the results for Equation (3.2).

The interaction between individuals who transition into spousal caregiving, $NC \rightarrow C$, and the period of treatment shows positive values for the physical measures. However, only the coefficient of self reported health is statistically significant. The results imply that transition into caregiving generates a slight improvement on how individuals rate their physical health. However, there seems to be no effect on actual physical capacities since the disability index shows no improvements. The depression coefficient is negative but not statistically significant.

The results are not enough to assess the long term effects, and should therefore be interpreted only as short run effects of providing care to a partner.

Table 3.7: Longitudinal Analysis - Three Differences Model

	Self Assessed Health (OProbit)	Depression (OProbit)	Disability Index (GLM)
main			
Treatment period	-0.005	0.041	0.021
	(0.033)	(0.032)	(0.014)
NC $\rightarrow$ C	0.034	0.037	0.047**
	(0.035)	(0.034)	(0.015)
NC $\rightarrow$ C in treatment period	0.078*	-0.015	0.015
	(0.039)	(0.039)	(0.018)

Regressions use a total of 7.704 individual surveys (from which 3,852 belong to the groups $NC \rightarrow C$) collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 17 countries. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. All individual and partner characteristics defined in Section 3.3.1 are used in the regressions but are omitted in the table.

To provide a glimpse into long run effects, I provide the same analysis for continued caregiving. I match continued caregivers on a first wave against "continued" non-caregivers. This approach is not as robust as the last one since, in this case, the caregiving status is less likely to be random. There must be a reason why people with similar health and similar partners have different decisions regarding caregiving. If those reasons are in any way related with the outcome variables, then results are biased. Consequently, conclusions should be taken with caution. Table 3.8 displays the results.

Table 3.8: Longitudinal Analysis - Three Differences Model

	Self Assessed Health (OProbit)	Depression (OProbit)	Disability Index (GLM)
Treatment period	-0.012	0.012	0.015
	(0.074)	(0.075)	(0.031)
C → C	0.252***	0.127	0.094**
	(0.076)	(0.073)	(0.034)
C → C in treatment period	-0.084	-0.020	-0.027
	(0.076)	(0.077)	(0.033)

Regressions use a total of 1.848 individual surveys (from which 924 belong to the group of continued caregivers, $C \rightarrow C$) collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 17 countries[8]. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. All individual and partner characteristics defined in Section 3.3.1 are used in the regressions but are omitted in the table.

Continued spousal caregivers present a negative variation in both measures of physical health and the measure of mental health. Nonetheless, the results are not statistically significant, indicating it is not possible to reject the null hypothesis. So, spousal caregivers and non-caregivers who have the same individual and partner characteristics in a given period, also display similar health outcomes one period later. The results depend on the belief that these matched caregivers and non-caregivers have the same health depreciation rates. It also depends on the belief that the estimations provide enough statistical power.[9]

Results using different pre-treatment matching calipers are in all identical to the ones presented in this sections, as shown in Table C.3 and Table C.4.

Summarizing, there are some slight positive short run effects of providing personal care to a partner in the way providers assess their physical health. Continued care does not appear to cause any

[9]Continued caregivers are few across observations which hinders the statistical power of these estimations.

differences in terms of health.

3.5.3 Other results

As in Section 3.4.4, the results of transitioning into partner caregiving may be reported by country. Those results are shown in Figure C.7, Figure C.8 and Figure C.9. The graphs show no heterogeneity in the results by countries, since countries are not statistically distinguishable between each other.

Similar to Section 3.4.5, this section may be extended to study the heterogeneity of the results with respect to gender and level of disability.

The results of the previous event study discriminated by gender are displayed in Figure C.10, Figure C.11 and Figure C.12. There is no evidence of observed heterogeneity in the partner caregiving effect across gender. Still, it is important to note that women present positive effects of becoming partner caregivers on the disability index.

The results of the previous event study discriminated by disability are displayed in Figure C.13, Figure C.14 and Figure C.15. There is no evidence of observed heterogeneity in the partner caregiving effect across disability intervals. On the disability index, however, the two higher intervals are the ones reporting gains of transitioning into partner caregiving.

3.6 Conclusion

The increase in life expectancy has a cost. If people live longer, they also live more time with frailer bodies. Consequently, authorities have increased their concerns about how to provide long term care to ageing individuals and any strategy should aim at optimizing between formal and informal care. But to do that, knowledge on the the pros and cons of each type of care is fundamental.

I use a very precise subsample of single living couples to provide a characterization of caregivers showing that, when compared to individuals with similar characteristics and equally sick partners, they present better health outcomes, which implies they constitute a positive selection. The results don't show any pattern when discriminated by countries. My approach also reinforces the importance of having into account correlations between partners' health which arise from family effects and shared household health investment decisions. Part of the result is causal, as I show that if transitions into caregiving are random, becoming a caregiver leads to gains in perceived physical health. However, there appear to be no gains in terms of mental health or actual physical health (level of disability). Continued caregiving does not show meaningful results. In SHARE, it is hard to observe partner caregivers across waves, which implies a small number of observations and, consequently weaker statistical power.

Positive outcomes for partner caregiving are good news, but they are only self-perceived. In this study I am looking at partners who provide personal care, which is only an entry level type of caregiving. The providers are mostly retired and perform simple tasks such as helping with getting out of bed, bathing or dressing. I do not distinguish between more severe cases or more demanding types of care. The results should therefore be interpreted in the respective context.

Bibliography

Anand, Krishnan S, M Fazıl Paç, and Senthil Veeraraghavan. 2011. "Quality–speed conundrum: Trade-offs in customer-intensive services." *Management Science*, 57(1): 40–56.

Bago d'Uva, Teresa, Owen O'Donnell, and Eddy Van Doorslaer. 2008. "Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans." *International journal of epidemiology*, 37(6): 1375–1383.

Bagust, Adrian, Michael Place, and John W Posnett. 1999. "Dynamics of bed use in accommodating emergency admissions: stochastic simulation model." *Bmj*, 319(7203): 155–158.

Bain, Christopher A, Peter G Taylor, Geoff McDonnell, and Andrew Georgiou. 2010. "Myths of ideal hospital occupancy." *Medical Journal of Australia*, 193(5): 311.

Baker, David W, Carl D Stevens, and Robert H Brook. 1991. "Patients who leave a public hospital emergency department without being seen by a physician: causes and consequences." *Jama*, 266(8): 1085–1090.

Bakx, Pieter, and Claudine de Meijer. 2013. "The influence of spouse ability to provide informal care on long-term care use." *Available at SSRN 2407926*.

Barbosa, Fátima, and Alice Delerue Matos. 2014. "Informal support in Portugal by individuals aged 50+." *European journal of ageing*, 11(4): 293–300.

Barros, P, S Machado, and J Simões. 2011. "Health system review." *Health systems in transition*, 13(4): 1–156.

Bekker, René, and Paulien M Koeleman. 2011. "Scheduling admissions and reducing variability in bed demand." *Health care management science*, 14(3): 237–249.

Benbassat, Jochanan, and Mark Taragin. 2000. "Hospital readmissions as a measure of quality of health care: advantages and limitations." *Archives of internal medicine*, 160(8): 1074–1081.

Bindman, Andrew B, Kevin Grumbach, Dennis Keane, Loren Rauch, and John M Luce. 1991. "Consequences of queuing for care at a public hospital emergency department." *Jama*, 266(8): 1091–1096.

Blom, Mathias C, Fredrik Jonsson, Mona Landin-Olsson, and Kjell Ivarsson. 2014*a*. "Associations between in-hospital bed occupancy and unplanned 72-h revisits to the emergency department: a register study." *International journal of emergency medicine*, 7(1): 1.

Blom, Mathias C, Fredrik Jonsson, Mona Landin-Olsson, and Kjell Ivarsson. 2014*b*. "The probability of patients being admitted from the emergency department is negatively correlated to in-hospital bed occupancy–a registry study." *International journal of emergency medicine*, 7(1): 1.

Blom, Mathias C, Karin Erwander, Lars Gustafsson, Mona Landin-Olsson, Fredrik Jonsson, and Kjell Ivarsson. 2015. "The probability of readmission within 30 days of hospital discharge is positively associated with inpatient bed occupancy at discharge–a retrospective cohort study." *BMC emergency medicine*, 15(1): 37.

Bobinac, Ana, N Job A Van Exel, Frans F H Rutten, and Werner B F Brouwer. 2010*a*. "Caring for and caring about : Disentangling the caregiver effect and the family effect." *Journal of Health Economics*, 29(4): 549–556.

Bobinac, Ana, N Job A Van Exel, Frans FH Rutten, and Werner BF Brouwer. 2010*b*. "Caring for and caring about: disentangling the caregiver effect and the family effect." *Journal of health economics*, 29(4): 549–556.

Boden, DG, A Agarwal, T Hussain, SJ Martin, N Radford, MS Riyat, K So, Y Su, A Turvey, and CI Whale. 2015. "Lowering levels of bed occupancy is associated with decreased inhospital mortality and improved performance on the 4-hour target in a UK District General Hospital." *Emergency Medicine Journal*, emermed–2014.

Bolin, K, B Lindgren, and P Lundborg. 2008. "INFORMAL AND FORMAL CARE AMONG SINGLE-LIVING ELDERLY IN EUROPE." 409(September 2007): 393–409.

Bonsang, Eric. 2009. "Does informal care from children to their elderly parents substitute for formal care in Europe ?" 28: 143–154.

Boulding, William, Seth W Glickman, Matthew P Manary, Kevin A Schulman, and Richard Staelin. 2011. "Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days." *The American journal of managed care*, 17(1): 41–48.

Brouwer, W. B.F., N. J.A. Van Exel, B. Van Gorp, and W. K. Redekop. 2006. "The CarerQol instrument: A new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations." *Quality of Life Research*, 15(6): 1005–1021.

Brouwer, Werner B. F., N. Job A. van Exel, Bernard van de Berg, Huib J. Dinant, Marc A. Koopmanschap, and Geertrudis A. M. van den Bos. 2004. "Burden of caregiving: Evidence of objective burden, subjective burden, and quality of life impacts on informal caregivers of patients with rheumatoid arthritis." *Arthritis Care & Research*, 51(4): 570–577.

Brouwer, Werner BF, N Job A van Exel, Marc A Koopmanschap, and Frans FH Rutten. 1999. "The valuation of informal care in economic appraisal: a consideration of individual choice and societal costs of time." *International Journal of Technology Assessment in Health Care*, 15(1): 147–160.

Chan, David C. 2016. "Teamwork and moral hazard: evidence from the emergency department." *Journal of Political Economy*, 124(3): 734–770.

Chandra, Amitabh, David Cutler, and Zirui Song. 2011. "Who ordered that? The economics of treatment choices in medical care." In *Handbook of health economics*. Vol. 2, 397–432. Elsevier.

Chen, Lena M, Wildon R Farwell, and Ashish K Jha. 2009. "Primary care visit duration and quality: does good care take longer?" *Archives of internal medicine*, 169(20): 1866–1872.

Chrusch, Carla A, Kendiss P Olafson, Patricia M McMillan, Daniel E Roberts, and Perry R Gray. 2009. "High occupancy increases the risk of early death or readmission after transfer from intensive care." *Critical care medicine*, 37(10): 2753–2758.

Clark, Derek, and Jan Abel Olsen. 1994. "Agency in health care with an endogenous budget constraint." *Journal of health economics*, 13(2): 231–251.

Coe, Norma B, Courtney Harold, and V A N Houtven. 2009. "CARING FOR MOM AND NEGLECTING YOURSELF ? THE HEALTH EFFECTS OF CARING FOR AN ELDERLY PARENT 1." 1010(July): 991–1010.

Cutler, David M. 1993. "The incidence of adverse medical outcomes under prospective payments." National Bureau of Economic Research.

Do, Young Kyung, Edward C Norton, Sally C Stearns, and Courtney Harold Van Houtven. 2015. "Informal care and caregiver's health." *Health economics*, 24(2): 224–237.

Duarte, Fabian. 2012. "Price elasticity of expenditure across health care services." *Journal of health economics*, 31(6): 824–841.

Dugdale, David C, Ronald Epstein, and Steven Z Pantilat. 1999. "Time and the patient–physician relationship." *Journal of general internal medicine*, 14(Suppl 1): S34.

Durbin, Janet, Elizabeth Lin, Crystal Layne, and Moira Teed. 2007. "Is readmission a valid indicator of the quality of inpatient psychiatric care?" *The journal of behavioral health services & research*, 34(2): 137–150.

Eichner, Matthew J. 1998. "The demand for medical care: What people pay does matter." *The*

American Economic Review, 88(2): 117–121.

Ellis, Randall P, and Thomas G McGuire. 1986. "Provider behavior under prospective reimbursement: Cost sharing and supply." *Journal of health economics*, 5(2): 129–151.

Ellis, Randall P, Bruno Martins, and Wenjia Zhu. 2017. "Health care demand elasticities by type of service." *Journal of health economics*, 55: 232–243.

Epstein, Stephen K, David S Huckins, Shan W Liu, Daniel J Pallin, Ashley F Sullivan, Robert I Lipton, and Carlos A Camargo Jr. 2012. "Emergency department crowding and risk of preventable medical errors." *Internal and emergency medicine*, 7(2): 173–180.

Forster, Alan J, Ian Stiell, George Wells, Alexander J Lee, and Carl Van Walraven. 2003. "The effect of hospital occupancy on emergency department length of stay and patient disposition." *Academic Emergency Medicine*, 10(2): 127–133.

Gattinoni, Luciano, Danilo Radrizzani, Bruno Simini, Guido Bertolini, Luca Ferla, Giovanni Mistraletti, Francesca Porta, Dinis R Miranda, et al. 2004. "Volume of activity and occupancy rate in intensive care units. Association with mortality." *Intensive care medicine*, 30(2): 290–297.

Godager, Geir, and Daniel Wiesen. 2013. "Profit or patients' health benefit? Exploring the heterogeneity in physician altruism." *Journal of health economics*, 32(6): 1105–1116.

Gorunescu, Florin, Sally I McClean, and Peter H Millard. 2002. "A queueing model for bed-occupancy management and planning of hospitals." *Journal of the Operational Research Society*, 53(1): 19–24.

Green, Linda V, and Vien Nguyen. 2001. "Strategies for cutting hospital beds: the impact on patient service." *Health services research*, 36(2): 421.

Grossman, Michael. 1972. "On the concept of health capital and the demand for health." *Journal of Political economy*, 80(2): 223–255.

Gruber, Jonathan, Thomas P Hoe, and George Stoye. 2018. "Saving lives by tying hands: The unexpected effects of constraining health care providers." National Bureau of Economic Research.

Heger, Dörte. 2017. "The mental health of children providing care to their elderly parent." *Health economics*, 26(12): 1617–1629.

Hiel, Laura, Mariëlle A Beenackers, Carry M Renders, Suzan JW Robroek, Alex Burdorf, and Simone Croezen. 2015. "Providing personal informal care to older European adults: Should we care about the caregivers' health?" *Preventive medicine*, 70: 64–68.

Hoot, Nathan R, and Dominik Aronsky. 2008. "Systematic review of emergency department crowding: causes, effects, and solutions." *Annals of emergency medicine*, 52(2): 126–136.

Kolstad, Jonathan T. 2013. "Information and quality when motivation is intrinsic: Evidence from surgeon report cards." *American Economic Review*, 103(7): 2875–2910.

Kowalski, Amanda. 2016. "Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care." *Journal of Business & Economic Statistics*, 34(1): 107–117.

Llácer, A, M V Zunzunegui, F Béland, and S H Zarit. 2002. "Correlates of wellbeing of spousal and children carers of disabled people over 65 in Spain." 3–9.

Martins, Bruno, and Luís Filipe. 2020. "Doctors' response to queues: Evidence from a Portuguese emergency department." *Health Economics*, 29(2): 123–137.

McCarthy, Melissa L, Scott L Zeger, Ru Ding, Scott R Levin, Jeffrey S Desmond, Jennifer Lee, and Dominik Aronsky. 2009. "Crowding delays treatment and lengthens emergency department length of stay, even among high-acuity patients." *Annals of emergency medicine*, 54(4): 492–503.

McClellan, Mark. 2011. "Reforming payments to healthcare providers: The key to slowing healthcare cost growth while improving quality?" *Journal of Economic Perspectives*, 25(2): 69–

92.

McCusker, Jane, Alain Vadeboncoeur, Jean-Frédéric Lévesque, Antonio Ciampi, and Eric Belzile. 2014. "Increases in emergency department occupancy are associated with adverse 30-day outcomes." *Academic Emergency Medicine*, 21(10): 1092–1100.

Nihtilä, Elina, and Pekka Martikainen. 2008. "Institutionalization of older adults after the death of a spouse." *American journal of public health*, 98(7): 1228–1234.

Ogden, Jane, Kheelna Bavalia, Matthew Bull, Stuart Frankum, Chris Goldie, Micaela Gosslau, Azita Jones, Sonia Kumar, and Kishor Vasant. 2004. ""I want more time with my doctor": a quantitative study of time and the consultation." *Family Practice*, 21(5): 479–483.

Pinquart, Martin, and Silvia Sörensen. 2003. "Differences between caregivers and noncaregivers in psychological health and physical health: A meta-analysis." *Psychology and Aging*, 18(2): 250–267.

Prince, Martin J, Aartjan TF Beekman, Dorly JH Deeg, Rebecca Fuhrer, S-L Kivela, Brian A Lawlor, Antonio Lobo, Hallgrimur Magnusson, Ingeborg Meller, Herman Van Oyen, et al. 1999. "Depression symptoms in late life assessed using the EURO–D scale: Effect of age, gender and marital status in 14 European centres." *The British Journal of Psychiatry*, 174(4): 339–345.

Riganti, Andrea, Luigi Siciliani, and Carlo V Fiorio. 2017. "The effect of waiting times on demand and supply for elective surgery: Evidence from Italy." *Health economics*, 26: 92–105.

Schmitz, Hendrik, and Matthias Westphal. 2015. "Short-and medium-term effects of informal care provision on female caregivers' health." *Journal of health economics*, 42: 174–185.

Schulz, Richard, and Scott R. Beach. 1999. "Caregiving as a risk factor for mortality: The caregiver health effects study." *Journal of the American Medical Association*, 282(23): 2215–2219.

Sewitch, Maida J, Jane McCusker, Nandini Dendukuri, and Mark J Yaffe. 2004. "Depression in frail elders: impact on family caregivers." *International Journal of Geriatric Psychiatry*, 19(7): 655–665.

Siciliani, Luigi, Anderson Stanciole, and Rowena Jacobs. 2009. "Do waiting times reduce hospital costs?" *Journal of Health Economics*, 28(4): 771–780.

Silver, David. 2016. "Haste or waste? Peer pressure and the distribution of marginal returns to health care." *UC Berkeley Department of Economics Unpublished Working Paper*.

Sivey, Peter. 2018. "Should I stay or should I go? Hospital emergency department waiting times and demand." *Health economics*, 27(3): e30–e42.

Swami, Megha, Hugh Gravelle, Anthony Scott, and Jenny Williams. 2018. "Hours worked by general practitioners and waiting times for primary care." *Health economics*, 27(10): 1513–1532.

Thomas, J William, and James J Holloway. 1991. "Investigating early readmission as an indicator for quality of care studies." *Medical care*, 29(4): 377–394.

Tucker, Anita L. 2004. "The impact of operational failures on hospital nurses and their patients." *Journal of Operations Management*, 22(2): 151–169.

Van Houtven, Courtney Harold, and Edward C Norton. 2004. "Informal care and health care use of older adults." *Journal of health economics*, 23(6): 1159–1180.

Vedhara, Kav, Nigel KM Cox, Gordon K Wilcock, Paula Perks, Moira Hunt, Stephen Anderson, Stafford L Lightman, and Nola M Shanks. 1999. "Chronic stress in elderly carers of dementia patients and antibody response to influenza vaccination." *The Lancet*, 353(9153): 627–631.

Whynes, David K. 1996. "Towards an Evidence-Based National Health Service?" *The Economic Journal*, 106(439): 1702–1712.

Yaffe, Kristine, Patrick Fox, Robert Newcomer, Laura Sands, Karla Lindquist, Kyle Dane, and Kenneth E Covinsky. 2002. "Patient and caregiver characteristics and nursing home placement in patients with dementia." *Jama*, 287(16): 2090–2097.

Zachariasse, Joany M, Nienke Seiger, Pleunie PM Rood, Claudio F Alves, Paulo Freitas, Frank J Smit, Gert R Roukema, and Henriette A Moll. 2017. "Validity of the Manchester Triage System in emergency care: A prospective observational study." *PloS one*, 12(2).

Zarit, Steven H, Karen E Reever, and Julie Bach-Peterson. 1980. "Relatives of the impaired elderly: correlates of feelings of burden." *The gerontologist*, 20(6): 649–655.

Appendix A

Chapter 1

A.1 Institutional Context

The Portuguese health care system is primarily characterized by a universal National Health Service (NHS) that is regulated at the federal level. A comprehensive description of the Portuguese health care system can be found in Barros, Machado and Simões (2011).

The NHS is designed such that each individual has an assigned gatekeeper (family doctor) that serves as the point of entry to obtain health care. However, patients often face delays between visiting the gatekeeper and obtaining access to specialty care. For this reason, it is common for patients to bypass the intended referral mechanism by going straight to the ED, where they can obtain all specialty care (including lab tests and exams) in a single day. Because of this, many visits to the emergency department are, in fact, not true medical emergencies.

Visiting the ED is not free to the patient. There is a small flat access charge that is meant to prevent excessive use of the ED and is not a sizable source of revenue. There is also a co-payment for any additional lab test, exam, or treatment procedure the patient receives, which are typically unknown to patients until the bill is received at the end of the visit. The access charge does not seem to be successful at preventing inappropriate use of the ED (Barros, Machado and Simões

2011). This is in line with recent research showing that individuals are not price sensitive with respect to emergency visits (see Duarte 2012 or Ellis, Martins and Zhu 2017).

Between 2011 and 2012 (our sample period) the billing system changed. In 2011 the access charge was 9.60 €. The OOP was a direct mapping from each specific service to a co-payment for hundreds of types of services (for example, the OOP for an X-Ray was 1.80 € and for an Magnetic Resonance Imaging procedure was 21.50 €).

In 2012 the access charge more than doubled, increasing to 20 €. For the diagnosis and treatment procedures, the OOP was defined as a piecewise mapping between the price of the service and a co-payment (e.g., procedures that cost between 60 € and 64.99 € have an OOP of 12 €). This means that in 2012 more expensive procedures had a weakly higher co-payment.

A.2 Robustness checks

A.2.1 Exogeneity of Visits to the ED

The institutional context of our setting suggests that since patients bypass the gate-keeping process and choose to go to the ED instead, arrivals to the ED might not be entirely exogenous, at least for those with low levels of urgency. The endogeneity of low-urgency visits has been documented by Sivey (2018), who shows that these patients are more likely to walk out of the ED without being seen by a doctor. To protect against this we test an alternative IV, in a fashion similar to that in Sivey (2018). We still use the number of patients that arrived in the previous 60 minutes, but only those that present levels of urgency ranging from yellow to red. Indeed, by doing this we remove the two lowest levels of urgency (blue and green)[1], whose arrivals to the ED are more likely to be a choice and bias our estimates.

In addition to including only the three most urgent arrivals as the basis for our instrumental vari-

[1]Ideally we would like to exclude yellow as well in order to have an even stronger robustness test, but the lack of variation when considering only orange and red and its incapacity to predict the size of the overall queue erode the strength of the instrument.

able, we use them as three separate instruments, rather than their sum. We choose to estimate this over-identified model in order to conduct over-identification tests on the validity of the instruments.

In Table A.4, column (1), the queue coefficient indicates that one more patient in the waiting room causes a decrease in length of stay of 2.7%. The value is similar to the one obtained in the standard specification, attesting to the robustness of the first set of results. In column (2), the coefficient states that one more person in the queue decreases the probability of being sent for further care or diagnosis by 0.59 percentage points. In column (3), the intensive margin still shows a negative, non-statistically significant, coefficient, with one more patient waiting decreasing the OOP by 0.26%. Again, the result is similar to the one in the standard specification. These results suggest that using all arrivals as the basis for the instrument does not lead to biased estimates.

Finally, Table A.4 shows, in the last two rows, the Hansen J statistic and corresponding p-values of the over-identification test. The p-value of 0.2 and 0.3 points indicate a low correlation between the different instruments and the predicted residuals, suggesting that our instruments satisfy the exclusion restriction.

A.2.2 Instrument Time Interval

The instrumental variable used in this paper was defined as the number of arrivals in the previous 60 minutes to the patient's arrival to the ED. Theoretically, the exclusion restriction should hold for other time intervals, as long as the episodes that require emergency care are random, after controlling for hour-specific fixed effects. The only constraint in selecting the appropriate interval becomes the relevancy of the instrument in predicting the size of the queue.

We re-estimate Equation (1.1) using different time intervals for the instrument, corresponding to periods of 15, 30, and 90 minutes and present our results in Table A.5. As we decrease this time interval, the F-test of the first stage regression increases, as do the point estimates for the queue variable. However, the estimated coefficients are not statistically different from those in the benchmark specification. Even though the 15 and 30 minutes IVs present higher first-stage F-statistics than the benchmark model, they do not perform consistently better across all of the model specifications used in this paper. The strength of these alternative IVs falls below 10 when estimating the model specification of Equation (1.2).

A.2.3 Night and Day

The results shown in Section 1.4.2 are average effects. Even though the model relies on hourly fixed effects for identification, it does not look for potential heterogeneity within each day. This is especially important since it is the presence of capacity constraints that forces doctors to change their behavior. Given that the number of arrivals to the ED is much higher during the day, it is likely that constraints bind during this time, which identifies our coefficients.

In this section we divide the analysis into two different spells, night and day, in order to understand whether the results are general across the entire day or if they are being driven by the particularities of one of those periods. We define night as the period from 10 p.m. to 8 a.m., which is associated with a lower trend in terms of arrivals to the ED (see Figure 1.4).

The results in Table A.6, in the appendix, show that during day periods the LOS coefficient is significant and roughly the same as the one from Section 1.4.2. The night coefficient is also negative, but higher in absolute values and not statistically significant. In the out-of-pocket analysis the coefficients tell a similar story. The extensive margin is important only during the day, again with a similar coefficient to the one in Section 1.4.2. The intensive margin is also consistent, during the day period. However, in the night period it shows greater sensitivity to the queue level, presenting a higher and statistically significant coefficient.

These results show that the overall estimates are being driven by day-period patients. The number of patients during the night period is rather small, and so there is not enough variation to identify the coefficient with precision.

A.2.4 Age Groups

Age is another potential factor of heterogeneity. In our main specification we included the variable age in a linear framework, promoting the simplicity of the model and believing it would be effective enough. In this section we want to provide more depth to the analysis of the age variable by dividing it into five-years age groups. We therefore run Equation (1.1) with age replaced by the abovementioned age groups. The age groups' coefficients are displayed in Figure A.4, in annex.

Panel A of Figure A.4 shows that LOS increases with the age of the patient in what seems to be a trend close to linear, thus corroborating our first and simpler approach. Panel B shows a similar pattern regarding out-of-pocket payments. Therefore, we can say that both time and use of resources increase with age of the patient.

The queue coefficient, not shown in the figures, is robust to this change, for each of the outcomes, showing the same values as in the main approach.

Table A.1: Statistics of Arrivals within a day

	Hour Max	Hour Min	Max/Min	Day/Night
Blue	10h	5h	21.06	5.52
Green	10h	5h	11.72	3.83
Yellow	15h	6h	5.88	2.52
Orange	11h	6h	4.23	2.05
Red	19h	7h	6.57	1.58

The two first columns show the hour of the day at which the maximum and minimum amount of arrivals occur, on average. The two last columns show the ratio between the maximum number of arrivals and the minimum number of arrivals, and the ratio between the number of arrivals during daytime and the number of arrivals during nighttime.

Table A.2: Robustness to different levels of fixed effects

	Benchmark	No FE	Y/M/D	Y/M/H	Y+M+D+H
Queue/10	-0.286***	-0.062***	-0.062***	-0.010	-0.011
	(0.061)	(0.023)	(0.018)	(0.012)	(0.011)
Blue	-1.297***	-1.274***	-1.229***	-1.218***	-1.217***
	(0.032)	(0.039)	(0.030)	(0.030)	(0.042)
Green	-0.886***	-0.877***	-0.858***	-0.852***	-0.852***
	(0.013)	(0.018)	(0.007)	(0.007)	(0.017)
Orange	0.652***	0.663***	0.654***	0.657***	0.656***
	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)
Red	-0.368***	-0.378***	-0.377***	-0.381***	-0.378***
	(0.084)	(0.090)	(0.054)	(0.054)	(0.089)
Female	0.066***	0.067***	0.066***	0.072***	0.071***
	(0.017)	(0.017)	(0.015)	(0.014)	(0.017)
Age	0.015***	0.015***	0.015***	0.015***	0.015***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	-0.001**	-0.001**	-0.001***	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results for the effects of queues on the log of length of stay. Results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. The benchmark triage color is yellow. Column 1 includes fixed effects at the Year-Month-Day-Hour level (15,801 levels). Column 2 has no fixed effect. Column 3 includes fixed effects at the Year-Month-Day level (670 levels). Column 4 includes fixed effects at the Year-Month-Hours level (528 levels). Column 5 includes year, month, day and hour fixed effects, but no interactions. First stage SW F-stats: 30.89 (column 1); 30.13 (column 2); 961.72 (column 3); 1185.70 (column 4); 268.14 (column 5).

Table A.3: Regressions by Urgency Level

Dependent variable:	Log(LOS)			1 if OOP > access charge		
	Blu&Gre	Yellow	Oran&Red	Blu&Gre	Yellow	Oran&Red
Blue and Green Queue/10	-0.278**			-0.125***		
	(0.109)			(0.034)		
Yellow Queue/10		-2.139**			-0.461**	
		(0.935)			(0.196)	
Orange and Red Queue/10			-13.659			-2.375
			(16.752)			(3.983)
Female	0.112***	-0.014	-0.071	0.012	-0.059***	0.022
	(0.020)	(0.017)	(0.092)	(0.009)	(0.009)	(0.038)
Age	0.014***	0.015***	0.016***	0.003***	0.003***	0.003***
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Female*Age	-0.001***	0.000	-0.000	0.001***	0.001***	-0.001
	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.001)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the effects of queues of some urgency level on the length of stay (first three columns) and out-of-pocket payment (last three columns) of patients belonging to that same urgency level. The data used in the regressions consider the arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. It divides into 154,000 blue and green arrivals, 101,829 yellow arrivals, and 12,971 orange and red arrivals. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level. First stage SW F-stats: 18.50 (columns 1 and 4); 19.16 (columns 2 and 5); 3.88 (columns 3 and 6).

Table A.4: Alternative Instrumental Variable

Dependent variable:	Log(LOS)	1 if OOP > access charge	Log(OOP|OOP > access charge)
Queue/10	-0.277***	-0.059**	-0.026
	(0.087)	(0.029)	(0.032)
Blue	-1.296***	-0.474***	-0.189***
	(0.034)	(0.009)	(0.012)
Green	-0.886***	-0.238***	-0.160***
	(0.012)	(0.004)	(0.004)
Orange	0.652***	0.179***	0.168***
	(0.012)	(0.003)	(0.005)
Red	-0.368***	-0.046***	0.466***
	(0.084)	(0.015)	(0.025)
Female	0.066***	-0.009	-0.027***
	(0.017)	(0.006)	(0.006)
Age	0.015***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	0.001***	-0.000*
	(0.000)	(0.000)	(0.000)
First-stage Instrument F-Stat	10.376	10.376	13.628
Hansen J-Statistic	3.072	0.204	0.421
Chi-sq(2) p-value	0.215	0.903	0.810

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the IV results for the effects of queue size on length of stay and out-of-pocket payments. This model uses previous arrivals of triage color red, orange, and yellow as the three IVs for queue. In LOS and OOP extensive margin regressions, results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. In OOP intensive margin regressions, results use a total of 73,609 arrivals, from Jan 2012-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level. The benchmark triage color is yellow.

Table A.5: Results with different IV time intervals

Dependent variable: Log(LOS) IV: Number of arrivals in the previous...	15 Minutes	30 Minutes	60 minutes (benchmark)	90 minutes
Queue/10	-0.341***	-0.297***	-0.286***	-0.259***
	(0.086)	(0.054)	(0.061)	(0.061)
Blue	-1.306***	-1.299***	-1.297***	-1.293***
	(0.036)	(0.033)	(0.032)	(0.032)
Green	-0.890***	-0.887***	-0.886***	-0.885***
	(0.015)	(0.012)	(0.013)	(0.013)
Orange	0.651***	0.652***	0.652***	0.653***
	(0.013)	(0.012)	(0.012)	(0.012)
Red	-0.368***	-0.368***	-0.368***	-0.368***
	(0.083)	(0.084)	(0.084)	(0.085)
Female	0.065***	0.066***	0.066***	0.067***
	(0.017)	(0.017)	(0.017)	(0.017)
Age	0.015***	0.015***	0.015***	0.015***
	(0.000)	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)
First-stage Instrument F-Stat	34.6	34.13	30.89	22.78

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The table shows the IV results for the effects of queue size on the treatment time for different IV specifications. Columns 1 to 4 display the coefficients when using as IV the number of persons arriving in the 15, 30, 60, and 90 minutes prior to the patient being called by a doctor. Results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level (15,801 levels). The benchmark triage color is yellow.

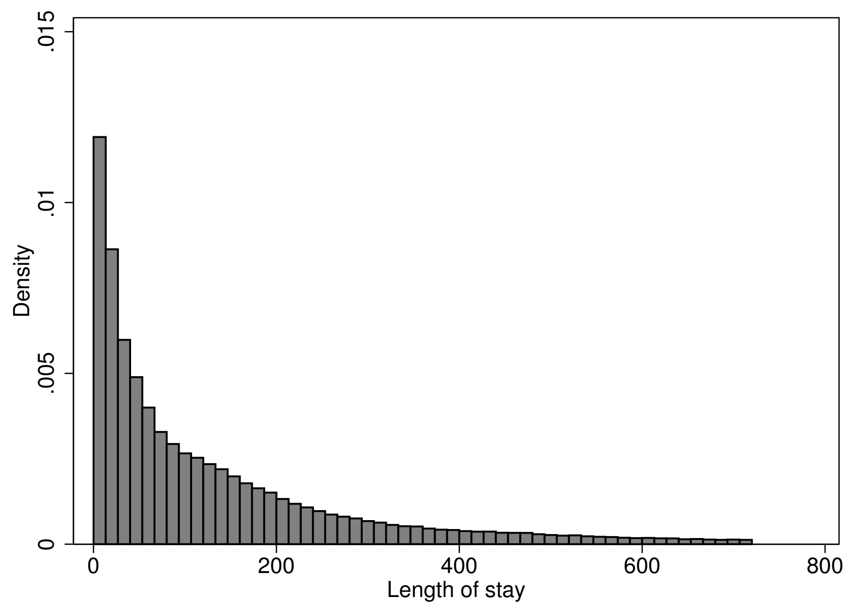
Table A.6: Day and Night regressions

Dependent variable:	Log(LOS)		1 if OOP > access charge		Log(OOP|OOP > access charge	
	Day	Night	Day	Night	Day	Night
Queue/10	-0.269***	-0.527	-0.088***	0.052	-0.013	-0.400**
	(0.055)	(0.381)	(0.016)	(0.111)	(0.018)	(0.174)
Blue	-1.333***	-1.066***	-0.485***	-0.411***	-0.182***	-0.308***
	(0.023)	(0.232)	(0.010)	(0.044)	(0.012)	(0.030)
Green	-0.897***	-0.860***	-0.243***	-0.205***	-0.160***	-0.200***
	(0.014)	(0.061)	(0.004)	(0.016)	(0.004)	(0.023)
Orange	0.647***	0.691***	0.174***	0.179***	0.170***	0.207***
	(0.012)	(0.058)	(0.004)	(0.014)	(0.005)	(0.040)
Red	-0.274***	-0.595***	-0.046**	-0.053**	0.448***	0.550***
	(0.095)	(0.134)	(0.018)	(0.027)	(0.028)	(0.049)
Female	0.059***	0.051	-0.007	-0.020*	-0.026***	-0.039***
	(0.019)	(0.038)	(0.007)	(0.011)	(0.006)	(0.014)
Age	0.015***	0.015***	0.003***	0.003***	0.003***	0.002***
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Female*Age	-0.001**	0.001	0.001***	0.001***	-0.000*	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
First-stage Instrument F-Stat	26.17	35.93	26.17	35.93	28.65	9.24

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

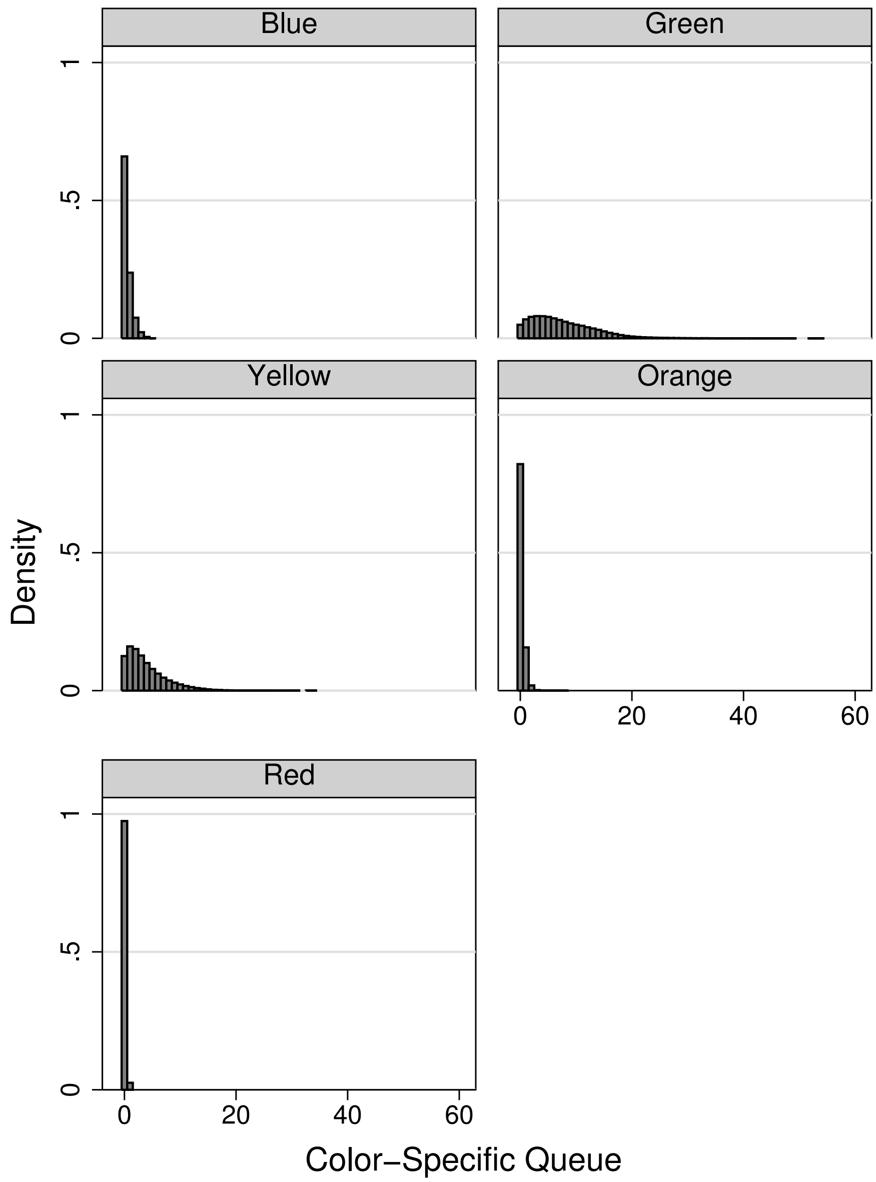
The table shows the IV results for the effects of queues on length of stay and out-of-pocket payments across night and day periods. For the LOS and OOP extensive margin regressions, results use a total of 276,061 arrivals in the emergency department of a Lisbon hospital, from Jan 2011-Oct 2012. For the OOP intensive margin regressions, results use a total of 73,609 arrivals, from Jan 2012-Oct 2012. Standard errors are cluster-robust at the hour level and shown in parentheses. All models include fixed effects at the Year-Month-Day-Hour level. The benchmark triage color is yellow.

Figure A.1: Length of Stay Distribution



Note: Length of stay, for each visit, is defined as the time (in minutes) between a patient being called to see a doctor for the first time and the time at which she is discharged by the doctor.

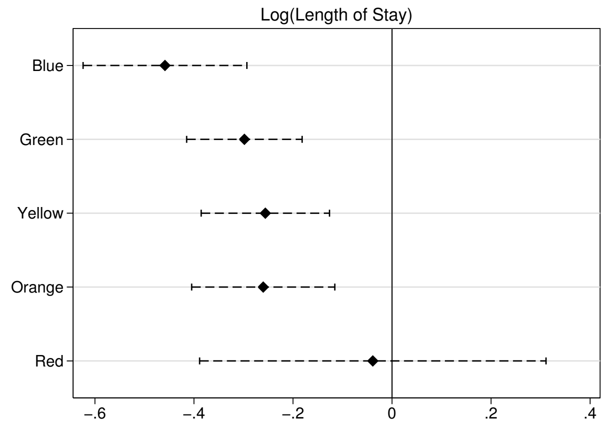
Figure A.2: Distribution of congestion by triage color of the visit



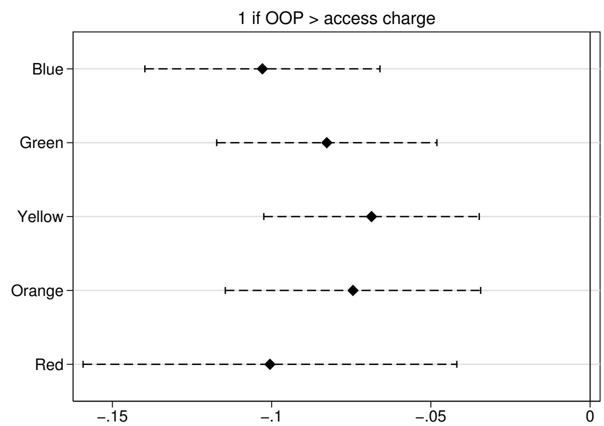
Note: Each panel subsets our sample according to the color that each visit was attributed at the triage. The color-specific queue is the number of people waiting that have the same triage color as the visit. For example, the "blue" panel uses all visits that had a "blue" triage level and the histogram shows the number of people waiting with a "blue" visit as well.

Figure A.3: Triage-specific slopes

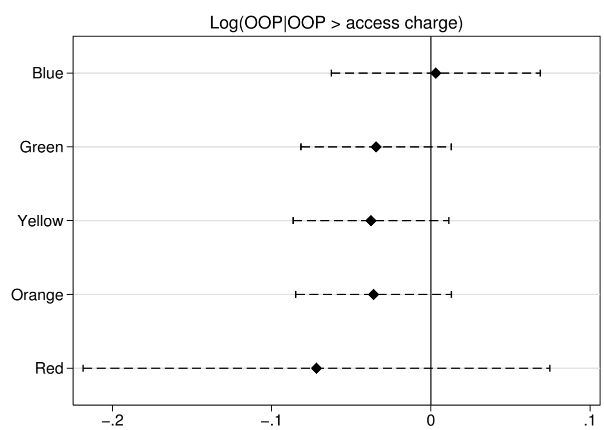
Panel A



Panel B



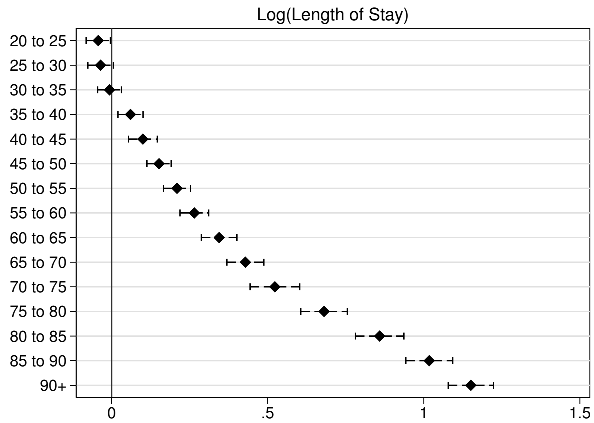
Panel B



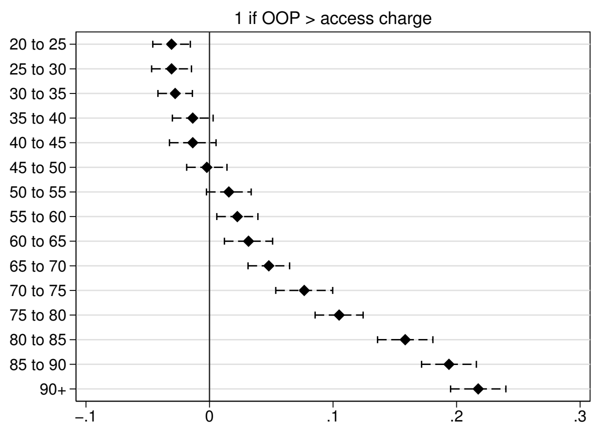
Note: Each figure represents the coefficients of the level of triage interacted with the queue in a regression model that includes Year-Month-Day-Hour fixed effects and uses the number of visits in the previous 60 minutes as the instrument for the queue. The second figure of Panel B includes data from 2012 only. Dashed lines at 95% confidence intervals.

Figure A.4: Age Groups specific slopes

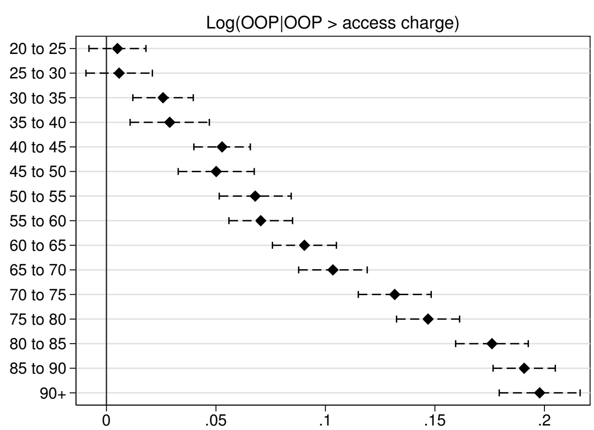
Panel A



Panel B



Panel B



Note: Each figure represents the coefficients of the age group of the patient (15 to 20 is the benchmark group) in a regression model that includes Year-Month-Day-Hour fixed effects and uses the number of visits in the previous 60 minutes as an instrument for queue size. The second figure of Panel B includes data from 2012 only.

Appendix B

Chapter 2

Table B.1: Main Estimates on the Probability of Readmission

Dependent variable: Readmission				
	Standard	Conservative	Standard	Conservative
Occupancy Two Days Before Discharge	0.035***	0.022***		
	(0.009)	(0.005)		
Turnover Rate t-2	-0.016***	-0.007***		
	(0.003)	(0.002)		
Occupancy Rate Variation t-2			0.076***	0.033***
			(0.008)	(0.007)
Male	0.008***	0.006***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Procedures	0.002***	0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.056	0.036	0.056	0.036

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of occupancy rates two days before discharge on the probability of a readmission episode. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

Table B.2: Main Estimates on the Probability of Readmission

Dependent variable: Readmission				
	Standard	Conservative	Standard	Conservative
Occupancy Three Days Before Discharge	-0.006	-0.000		
	(0.007)	(0.005)		
Turnover Rate t-3	-0.005	-0.003		
	(0.003)	(0.001)		
Occupancy Rate Variation t-3			0.022**	0.018***
			(0.007)	(0.004)
Male	0.008***	0.006***	0.008***	0.006***
	(0.000)	(0.000)	(0.000)	(0.000)
Number of Procedures	0.002***	0.000***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.056	0.036	0.056	0.036

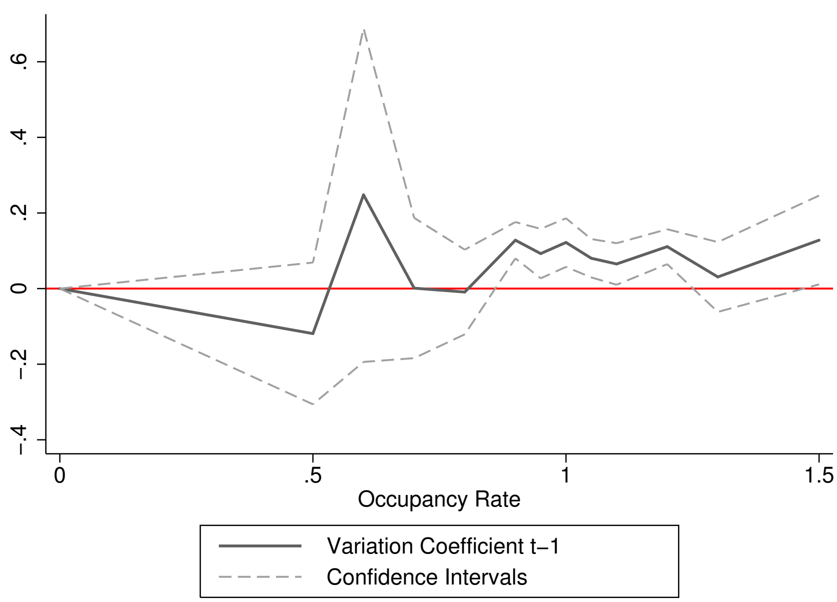
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the results of the effects of occupancy rates three days before discharge on the probability of a readmission episode. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

B.1 Interaction of variation and high occupancy

In Section 2.5.3 I answered a question about the non-linearity of the occupancy coefficient, but the same concern applies to occupancy variation. In the same line of thought, occupancy variation should only impact health outcomes if the hospital is already crowded. Again, having an increase of 10 percentage points in occupancy when the hospital is at 50% capacity should have little to no importance. But experiencing such increase in occupancy for a hospital which is already at 100% would likely cause problems. Theoretically it would be expected that a decision to discharge would be more sensible to variations when hospitals are crowded. To test this hypothesis, I use different intervals of occupancy and run individual regressions for each of those intervals. The results for the coefficients on occupancy variation one day before discharge are showed in Figure B.1. The corresponding conservative approach graph is displayed in Figure B.2.

Figure B.1: Variation in Occupancy by Occupancy Intervals



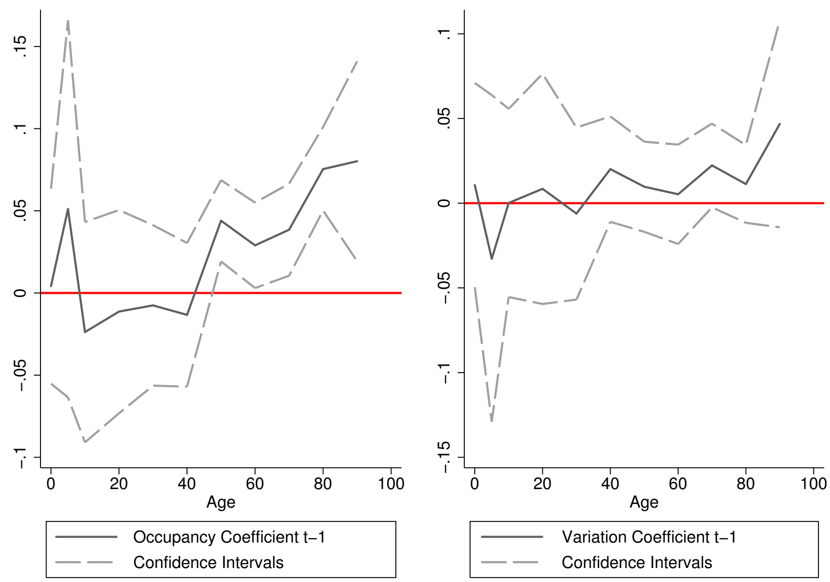
Note: This figure shows the effects of occupancy variation one day before patient's discharge on the likelihood of readmission by level of occupancy, using the standard approach.

Positive variation in occupancy rates one day before patient's discharge does not have any significant effect on the probability of readmission if occupancy is lower than 85%. From then on, positive variation in occupancy rates is associated with higher probability of readmission. For high enough values of occupancy, the probability increases by an average of 0.01 percentage points if occupancy variation increases by 10 percentage points. Physicians manage the hospital resources (including their own time), and when observing an increase in occupancy on an already-crowded hospital they may decide to force some discharges to accommodate for new patients.

Figure B.2: Variation in Occupancy by Occupancy Intervals (conservative)

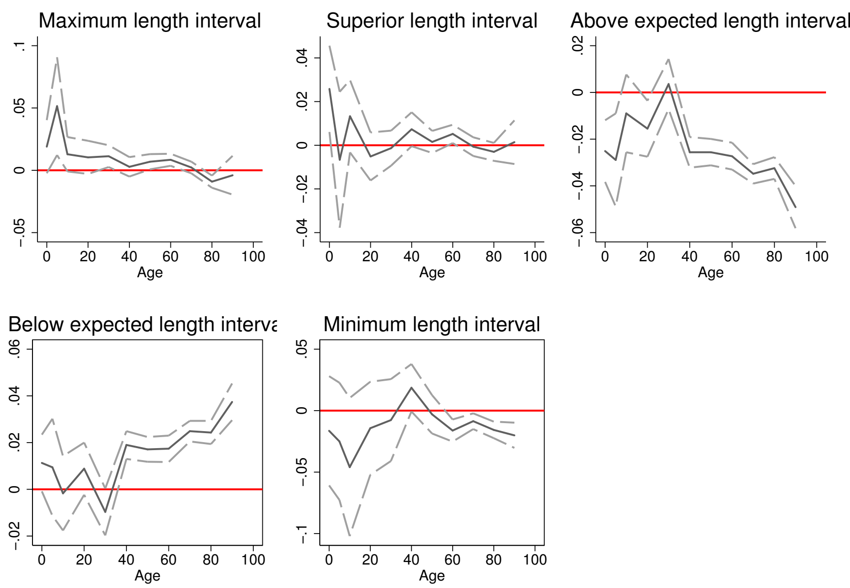
Note: This figure shows the effects of occupancy variation one day before patient's discharge on the likelihood of readmission by level of occupancy, using the conservative approach.

Figure B.3: Occupancy Coefficients by Age (conservative)



Note: This figure shows the effects of occupancy rate and occupancy variation one day before patient's discharge on the likelihood of readmission by age, using the conservative approach.

Figure B.4: Length of Stay Thresholds by Age (conservative)



Note: This figure shows the effects of occupancy rate one day before patient's discharge on the likelihood of readmission by thresholds of length of stay (for the respective DRG and severity level) and by age, using the conservative approach.

B.2 Occupancy, readmission and days of the week

The analysis I performed so far uses fixed effects at the week level, considering very fine variation in occupancy variables that occur in a given week, in a given hospital. Still, I do not consider days of the week, which are important determinants of hospital occupancy fluctuations mostly through the supply side of the hospital. Specifically, weekends play a big role in determining supply, as hospitals do not have as much resources available (physicians, nurses and other staff), which in turn, determine the level of occupancy of the hospital. Even though these effects must be taken into account, increasing the range of fixed effects eliminates too much variation in the model[1], not allowing the use of multiplicative fixed effects.

Given the difficulties of including days of the week in the full model, I evaluate them separately. Table B.3 shows the average values for occupancy variation rate, occupancy rate, standard readmission and conservative readmission. These descriptive statistics show important differences between working days and weekends. Not only weekends present much lower occupancy rates, they also present lower readmission rates. This alerts for the possibility of days of the week explaining part of the relationship between occupancy variables and readmissions.

Table B.4 displays the readmission rates for all days of the week when occupancy rates are above and below 85%[2]. Even conditioning on occupancy, days of the week still present strong differences between working days and weekends, implying that the weekend effect is there independently of occupancy rates. The discharging criteria switches across days of the week, and patients are usually discharged before the weekend, preparing the hospital for a period of lower resource availability. Rushed discharges are more likely to occur during working days rather than weekends.

[1]Including year/month/week/day-of-week/hospital fixed effects eliminates all variation in the model since occupancy rates are counted by hospital by day of discharge. Including year/month/day-of-week/hospital fixed effects also leave very little variation.

[2]This value was chosen based on Figure B.1, which shows that variation in occupancy only presents significant effects after occupancy reaches 85%.

Table B.3: Descriptive statistics on days of the week

	Variation Rate	Occupancy Rate	Standard Readmission	Conservative Readmission
Monday	0.67%	103.2%	9.4%	4.7%
Tuesday	0.07%	103.6%	9.4%	4.7%
Wednesday	-0.3%	103.8%	9.1%	4.6%
Thursday	-1.07%	102.9%	9.0%	4.5%
Friday	-5.65%	100.7%	9.6%	4.9%
Saturday	0.43%	0.91%	6.3%	3.1%
Sunday	3.37%	0.93%	5.8%	2.9%

The table shows the average values of variation in occupancy rate, occupancy rate, standard readmission rate and conservative readmission rate.

Table B.4: Descriptive statistics on days of the week

	Occupancy $>$ 85%		Occupancy $<$ 85%	
	Standard Readmission	Conservative Readmission	Standard Readmission	Conservative Readmission
Monday	9.5%	4.8%	7.5%	3.5%
Tuesday	9.5%	4.7%	7.5%	3.4%
Wednesday	9.2%	4.6%	6.9%	3.5%
Thursday	9.1%	4.6%	7.3%	3.5%
Friday	9.7%	4.9%	8.1%	4.2%
Saturday	6.4%	3.2%	5.8%	2.9%
Sunday	5.9%	2.9%	5.3%	2.7%

The table shows the average values of standard readmission rates and conservative readmission rates for two different intervals of occupancy rates, below 85% and above 85%.

Days of the week present relevant differences between each other, due to the structure of hospital supply. Thus it is worth asking whether occupancy variables may have different effects on readmissions across each day. To provide some insight on the matter, Table B.5 provides the coefficients of the occupancy variables one day before discharge for seven regressions, one for each day of the week.

Table B.5: Main Estimates on the Probability of Readmission per day of the week

Dependent variable: Readmission	Standard Occupancy t-1	Conservative Occupancy t-1	Standard Variation t-1	Conservative Variation t-1
Occupancy - Monday	0.015	0.011	0.021	0.023
	(0.010)	(0.007)	(0.026)	(0.021)
Occupancy - Tuesday	0.032***	0.025***	-0.021	-0.033*
	(0.008)	(0.007)	(0.023)	(0.017)
Occupancy - Wednesday	0.009	0.004	0.010	0.008
	(0.008)	(0.005)	(0.020)	(0.016)
Occupancy - Thursday	0.024**	0.016*	-0.023	-0.028
	(0.008)	(0.007)	(0.025)	(0.015)
Occupancy - Friday	0.023*	0.008	-0.008	-0.011
	(0.010)	(0.006)	(0.021)	(0.015)
Occupancy - Saturday	0.014	0.009	-0.006	-0.008
	(0.010)	(0.008)	(0.021)	(0.021)
Occupancy - Sunday	0.021	0.009	-0.020	-0.009
	(0.015)	(0.011)	(0.045)	(0.034)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the coefficients on occupancy and occupancy variation one day before discharge per day of the week. Each coefficient belongs to an independent equation ran only with discharges belonging to the respective day of the week. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Hospital level (123 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

The results show that higher occupancy rates in the day before discharge are positive for all days of the week, but Tuesday, Thursday and Friday (standard approach) present coefficients statistically different from zero. This means that when occupancy is higher on Mondays, Wednesdays and Thursdays (standard approach), patients discharged in the next day are more likely to be readmitted. Regarding variation in occupancy rates, the results are mostly negative and non-statistically significant.

B.2.1 Occupancy and Readmission per MDC

The analysis so far has used 7 Major Diagnosis Categories to determine the correlation between occupancy variables and readmissions. They were chosen among all MDCs because they present higher rates of readmission. Now, it is time to understand which MCDs are driving the results. For that purpose, I run again the regression in Equation (2.1), but by MDC. The results of the 7 regressions are displayed in table Table B.6.

Table B.6 shows that the respiratory system, circulatory system and hepatic system are the most relevant groups when explaining the correlation between occupancy variables and readmissions. They are the three groups that show higher coefficients across all specifications. They are robust across occupancy variables, being statistically significant both using occupancy one day before discharge, and variation in occupancy one day before discharge. They are also robust across the readmission methodology showing higher and statistically significant results for both approaches of counting, the standard and the conservative approach.

These three categories are likely to be the ones requiring more resources in earlier stages of hospitalizations. A person suffering from a stroke, for instance, needs a stronger intervention at arrival to avoid greater consequences, but after stabilization does not need as much medical attention. It is possible that when hospitals become crowded, physicians try to alleviate pressure by discharging stabilized patients. However, if they do it too soon, they run the risk of promoting readmissions.

Table B.6: Main Estimates on the Probability of Readmission per MDC

Dependent variable: Readmission	Standard Occupancy t-1	Conservative Occupancy t-1	Standard Variation t-1	Conservative Variation t-1
Occupancy - Nervous	0.065**	0.017	0.046	-0.008
	(0.020)	(0.012)	(0.026)	(0.013)
Occupancy - Respiratory	0.171***	0.095***	0.178***	0.088***
	(0.021)	(0.017)	(0.022)	(0.018)
Occupancy - Circulatory	0.070***	0.053***	0.069***	0.027*
	(0.017)	(0.012)	(0.017)	(0.012)
Occupancy - Digestive	0.030	0.007	0.039**	0.014
	(0.017)	(0.008)	(0.015)	(0.009)
Occupancy - Hepatic	0.098***	0.069**	0.150***	0.106***
	(0.027)	(0.023)	(0.030)	(0.021)
Occupancy - Musculoskeletal	0.011	-0.001	0.008	0.000
	(0.010)	(0.006)	(0.017)	(0.009)
Occupancy - Urinary	0.049*	-0.001	0.048	0.016
	(0.024)	(0.020)	(0.034)	(0.024)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the coefficients on occupancy and occupancy variation one day before discharge per MDC. Each coefficient belongs to an independent equation ran only with discharges belonging to the respective MDC. Results use a total of 1,687,348 discharges in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Month-Week-Hospital level (9,286 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

B.2.2 Occupancy and Readmission per months

Months are very different and have unique characteristics. Figure 2.1 exemplifies how occupancy rates vary throughout the year. The relationship between occupancy rates and readmissions may also change across months. In this section, I analyze this relation independently for each month. Table B.7 shows that occupancy rates one day before discharge are specially impactful in driving readmission episodes in the months of January, February, April and July, with positive statistically significant coefficients both in the standard and conservative approach. Occupancy variation one day before discharge affects readmissions mostly in January, March, April and June, when the coefficients are statistically significant in both the standard and the conservative approach. The results are stronger in the first half of the year which is typically a period of higher pressure due to seasonal diseases. As displayed in Figure 2.1, it is also the period with higher occupancy rates. Previously, in Section 2.5.3, I showed that the result is not linear and being discharged at higher levels of occupancy increases the likelihood of readmission.

B.2.3 Occupancy and Readmission by hospital

As described in Section 2.2, hospitals in the Portuguese Health Service may be very heterogeneous. They are located in different regions of the country, have different sizes and are organized in various different ways. In this section, I study the possibility of the effects of occupancy rates on readmissions being heterogeneous across the various hospital characteristics. Table B.8 divides all the hospitals in the dataset by region, type and size. ARS Norte, Centro, LVT, Alentejo and Algarve, stand for the regions of North, Center, Lisbon and the Tagus Valley, Alentejo and Algarve. A comprehensive map of the country is illustrated in Figure B.8. Hospital stands for single hospitals, Hospital Center is a group of hospitals and ULS is a Local Unit of Health, which integrates inpatient care, primary care and other services. The sizes are divided into Very big, Big, Medium and Small. I classify hospitals as very big if they belong to the fourth (upper) quartile in

Table B.7: Main Estimates on the Probability of Readmission per month

Dependent variable: Readmission	Standard Occupancy t-1	Conservative Occupancy t-1	Standard Variation t-1	Conservative Variation t-1
January	0.038***	0.019**	0.116***	0.058***
	(0.007)	(0.007)	(0.024)	(0.017)
February	0.042*	0.053***	0.067	0.026
	(0.021)	(0.006)	(0.034)	(0.028)
March	0.017	-0.001	0.100***	0.031*
	(0.016)	(0.007)	(0.018)	(0.015)
April	0.057***	0.039*	0.093***	0.032***
	(0.012)	(0.020)	(0.026)	(0.008)
May	0.032	0.018	0.077***	0.028
	(0.020)	(0.017)	(0.021)	(0.018)
June	0.053**	0.035	0.087***	0.032*
	(0.017)	(0.019)	(0.016)	(0.014)
July	0.065***	0.046***	0.015	0.024
	(0.008)	(0.012)	(0.030)	(0.015)
August	0.026	-0.002	0.080**	0.030
	(0.014)	(0.011)	(0.031)	(0.015)
September	0.012	0.010	0.008	0.009
	(0.013)	(0.010)	(0.013)	(0.010)
October	0.013	0.003	0.022	0.009
	(0.009)	(0.014)	(0.017)	(0.011)
November	0.026	0.022	0.060***	0.028
	(0.020)	(0.013)	(0.015)	(0.016)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the coefficients on occupancy and occupancy variation one day before discharge per month. Each coefficient belongs to an independent equation ran only with discharges belonging to the respective month. Results use a total of 1,687,348 discharges, separated throughout each month regression, in 41 Portuguese Public Hospitals, from 2014 to 2016. Standard errors are clustered at the week level and shown in parenthesis. All models include fixed effects at the Year-Hospital level (123 levels), age groups fixed effects (11) and DRG-Severity fixed effects (595).

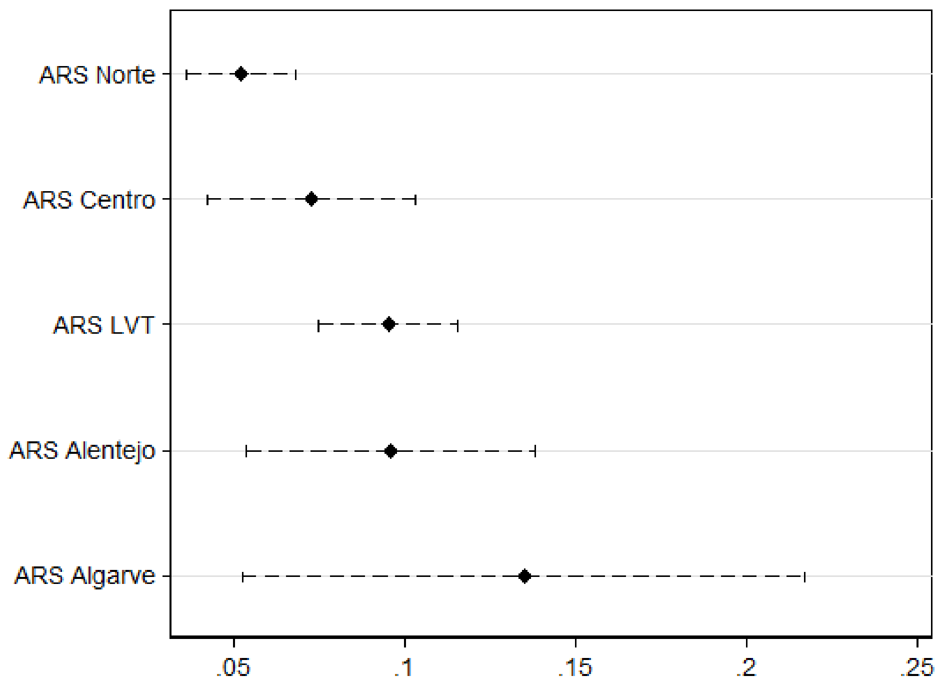
their percentage of hospitalizations. They are classified as big if they belong to the third quartile, medium if they belong to the second quartile, and small if they belong to the first quartile. Finally, I add an observation if hospitals are managed by private companies as a product of a partnership between the state and a private institution (PPP), or if they are university hospitals.

The results present almost no heterogeneity across regions. Figure B.5 shows the coefficients of the occupancy one day before discharge on readmissions, divided by region of the hospitals. All regions display positive and statistically significant coefficients, which shows that the effects found in Section 2.5.2 are true in all regions. Only ARS Norte and ARS LVT display significantly different coefficients from each other, with the southern region presenting a higher effect of occupancy rates one day before discharge on the probability of readmission.

Figure B.6 shows the occupancy coefficients differentiated by the type of hospital. All types show positive point estimates, statistically different from zero. PPP hospitals and university hospitals show lower coefficients, but not distinguishable from other hospital centers, hospitals and ULS. The impact of occupancy rates one day before discharge on readmission probability is not heterogeneous across hospital types.

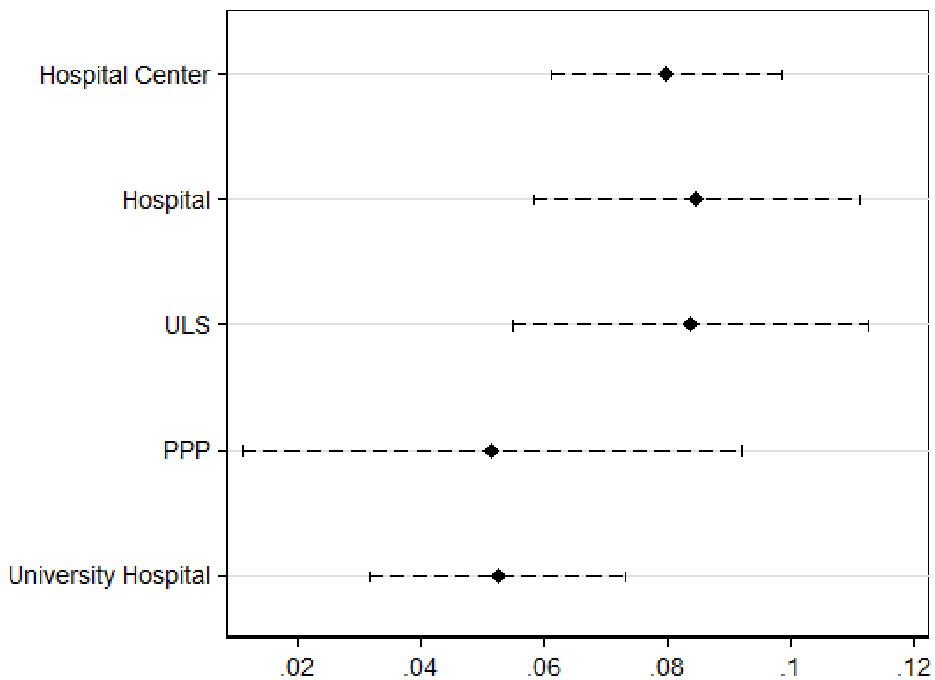
Figure B.7 shows the occupancy coefficients distinguished by size. All four sizes show positive and statistical significant coefficients, meaning that, independently of size, higher occupancy rates one day before discharge increase the probability of readmission. There is no heterogeneity between different-sized hospitals.

Figure B.5: Occupancy and readmissions by hospital region



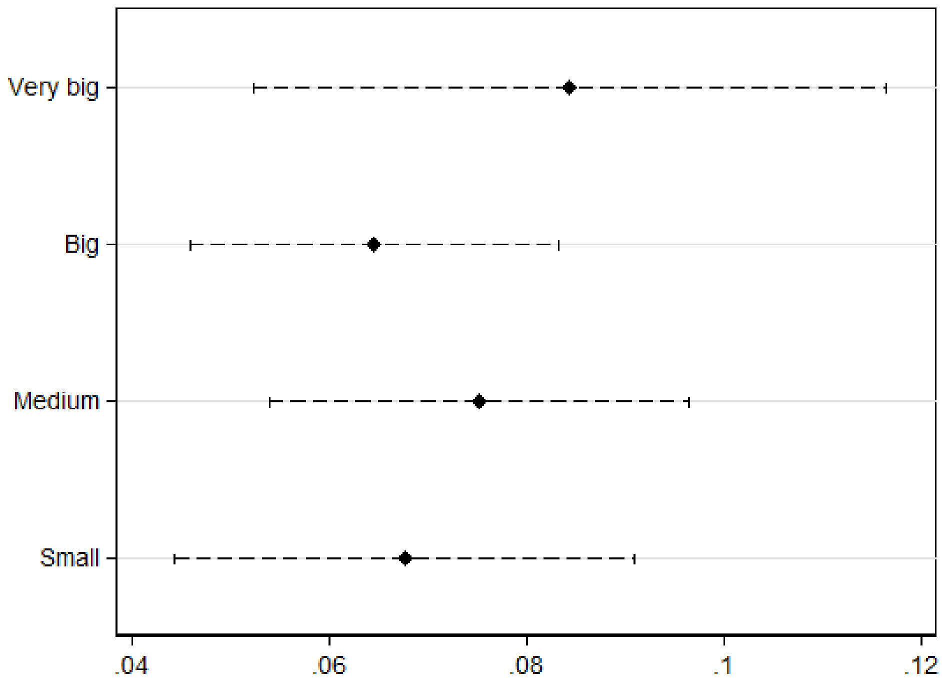
Note: The figure shows the effects of occupancy rates one day before patient's discharge on the likelihood of readmission discriminated by hospital region (North, Center, Lisbon and the Tagus Valley, Alentejo and Algarve), using the standard approach.

Figure B.6: Occupancy and readmissions by hospital type



Note: The figure shows the effects of occupancy rates one day before patient's discharge on the likelihood of readmission discriminated by hospital type (Hospital center, Hospital, ULS, PPP and University Hospital), using the standard approach.

Figure B.7: Occupancy and readmissions by hospital size



Note: The figure shows the effects of occupancy rates one day before patient's discharge on the likelihood of readmission discriminated by hospital size (Very big, Big, Medium and Small), using the standard approach.

Table B.8: List of hospitals

Hospital	Region	Type	Size	Observation
1	ARS LVT	Hospital	Medium	
2	ARS LVT	Hospital	Big	
3	ARS Norte	Hospital	Small	
4	ARS Norte	Hospital	Big	University Hospital
5	ARS LVT	Hospital	Small	PPP
6	ARS Centro	Hospital Center	Medium	
7	ARS Norte	Hospital Center	Medium	
8	ARS Algarve	Hospital Center	Big	University Hospital
9	ARS LVT	Hospital Center	Small	
10	ARS Centro	Hospital Center	Small	University Hospital
11	ARS Norte	Hospital Center	Medium	
12	ARS Norte	Hospital Center	Big	
13	ARS LVT	Hospital Center	Very big	University Hospital
14	ARS LVT	Hospital Center	Very big	University Hospital
15	ARS LVT	Hospital Center	Big	
16	ARS Centro	Hospital Center	Big	
17	ARS Norte	Hospital Center	Small	
18	ARS LVT	Hospital Center	Medium	
19	ARS LVT	Hospital Center	Small	
20	ARS Norte	Hospital Center	Big	University Hospital
21	ARS LVT	Hospital Center	Small	
22	ARS Norte	Hospital Center	Very big	University Hospital
23	ARS Norte	Hospital Center	Big	
24	ARS Centro	Hospital Center	Big	
25	ARS Centro	Hospital Center	Very big	University Hospital
26	ARS Norte	Hospital Center	Medium	
27	ARS Alentejo	Hospital	Small	
28	ARS Centro	Hospital	Small	
29	ARS LVT	Hospital	Medium	PPP
30	ARS Norte	Hospital	Small	
31	ARS Centro	Hospital	Small	
32	ARS Norte	Hospital	Small	
33	ARS LVT	Hospital	Medium	
34	ARS Norte	ULS	Medium	
35	ARS Alentejo	ULS	Small	
36	ARS Centro	ULS	Small	
37	ARS Alentejo	ULS	Small	
38	ARS Centro	ULS	Small	
39	ARS Alentejo	ULS	Small	
40	ARS Norte	ULS	Small	
41	ARS LVT	Hospital	Small	PPP

The table shows all the hospitals in the dataset by region type and size.

Figure B.8: Map of ARS

Source: Institutional site of ACSS.

Appendix C

Chapter 3

Table C.1: Summary Statistics

	Mean	Min	Max
Self Assessed Health	2.92	1 (very bad)	5 (very good)
Depression Scale	9.80	0 (very depressed)	12 (not depressed)
Disability Index	0.83	0.06 (very limited)	0.98 (not limited)
Education	10.93	0	25
Age	64.40	24	99
Gender (male)	49.99%	0%	100%
Partner Caregiver	5.24%	0%	100%
Employed	22.66%	0	100%
Retired	61.71%	0	100%
Homemaker	9.55%	0	100%
Sick or disabled	2.80%	0	100%
Make Ends Meet	2.98	1 (with great difficulty)	4 (easily)
Observations	99,613		

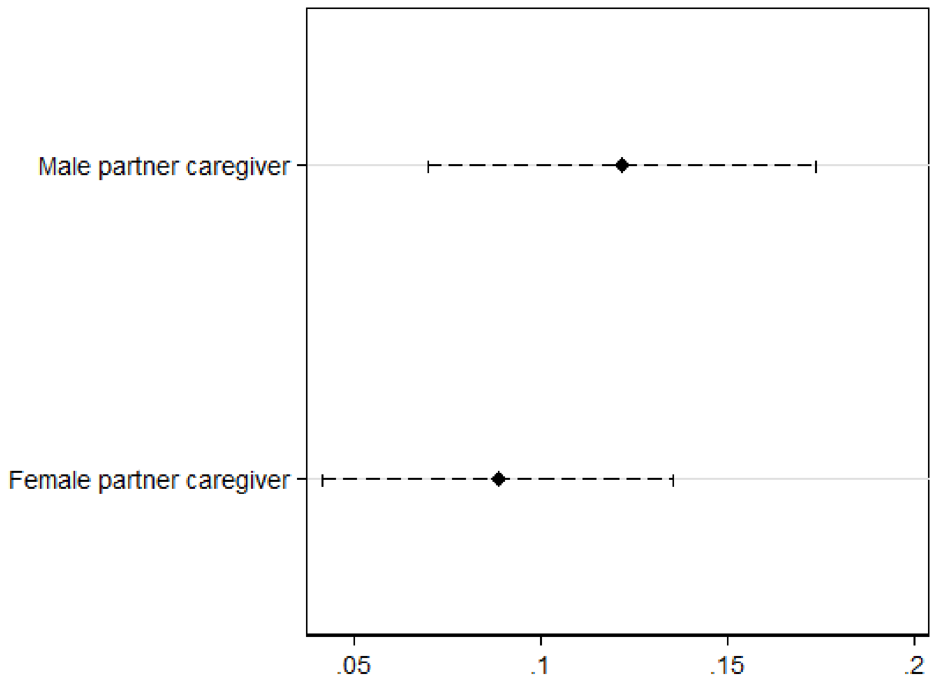
The table displays the overall summary statistics for the sample used in this study. The first column shows the averages for each variable, while the two last columns show the maximum and minimum values, respectively.

Table C.2: Treated vs. Matched

Means/Shares of treated and matched variables	Treated	Matched
Parter's Self Assessed Health	2.03	2.05
Partner's Depression Scale	8.26	8.25
Partner's Disability Index	0.64	0.65
Partner's Age	70.67	70.81
Gender	0.46	0.48
Age	69.87	70.22
Make Ends Meet	2.70	2.70
Retired	0.72	0.71
Employed or Self-Employed	0.11	0.11
Unemployed	0.02	0.02
Permanently sick or disabled	0.04	0.04
Homemaker	0.11	0.11
Other Status	0.01	0.01

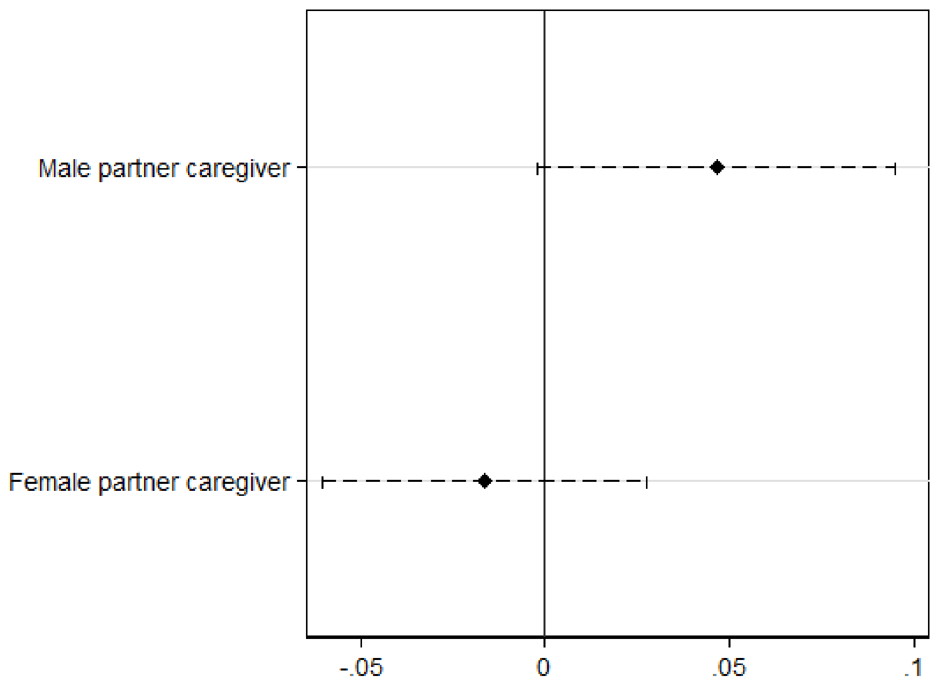
The table shows the averages of the variables used in the matching process. The factor variables of employment and gender display shares. Each group, matched and treated, amount to 5,082 individuals each.

Figure C.1: Partner caregiving and SAH by gender



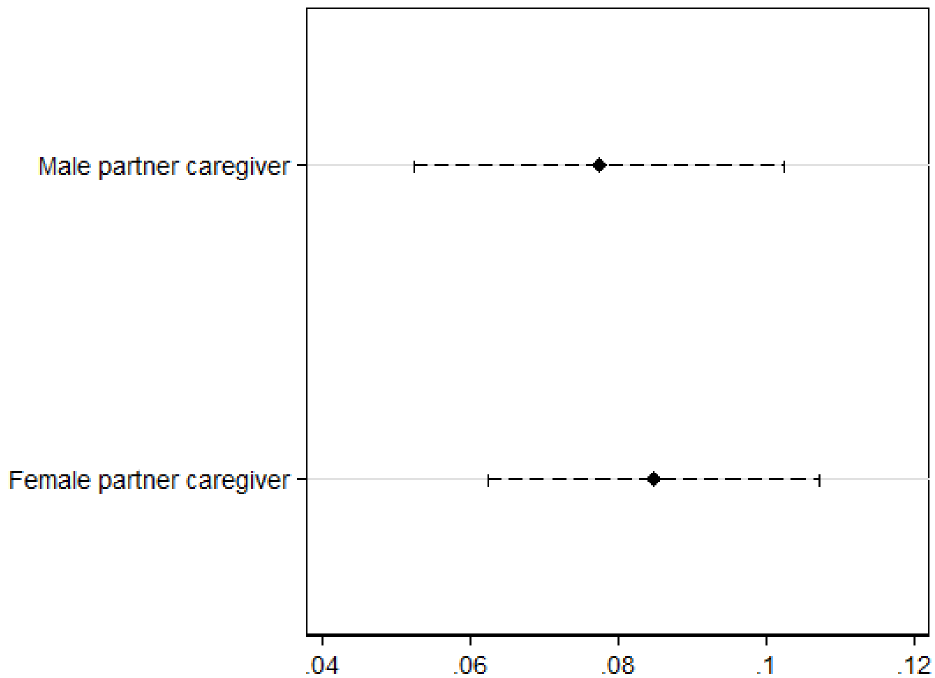
The graph shows the coefficients of the variable "Partner caregiver" discriminated by gender, in the SAH regression. The model includes both individuals and partner characteristics.

Figure C.2: Partner caregiving and depression by gender



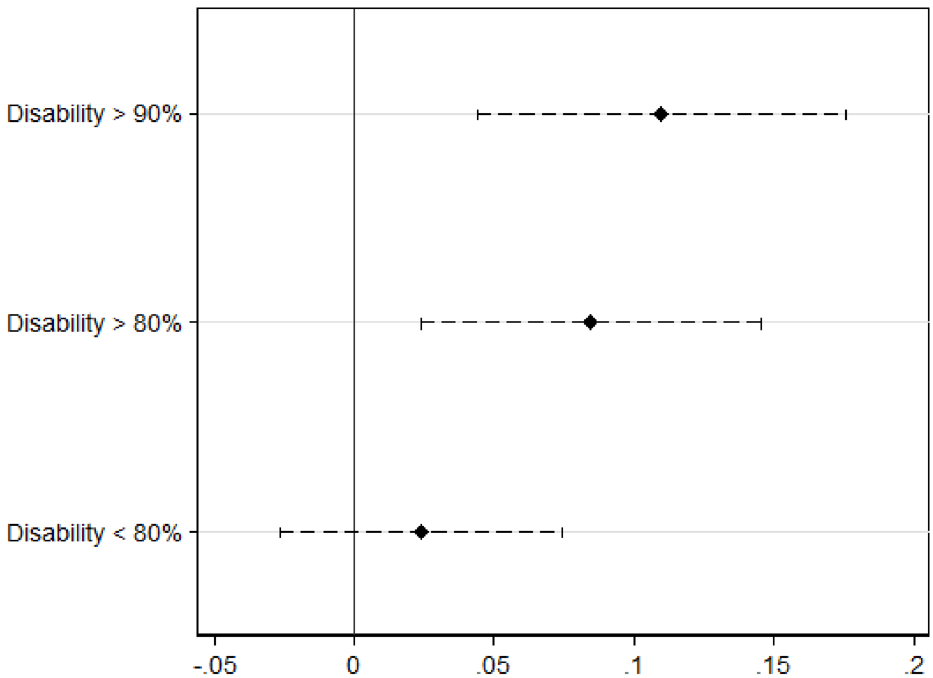
The graph shows the coefficients of the variable “Partner caregiver” discriminated by gender, in the depression regression. The model includes both individuals and partner characteristics.

Figure C.3: Partner caregiving and disability by gender



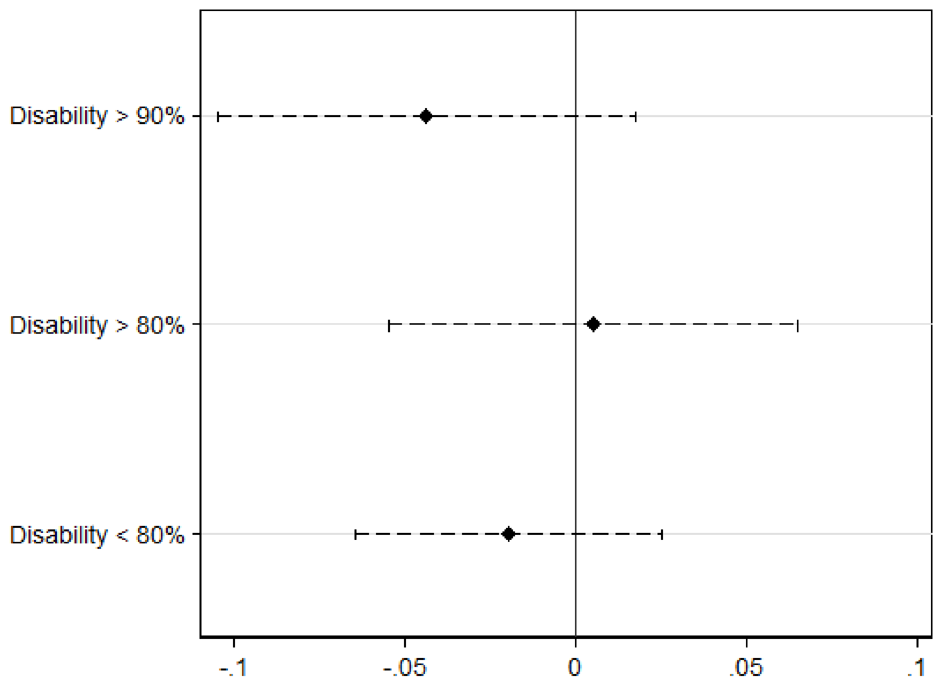
The graph shows the coefficients of the variable “Partner caregiver” discriminated by gender, in the disability regression. The model includes both individuals and partner characteristics.

Figure C.4: Partner caregiving and SAH by disability



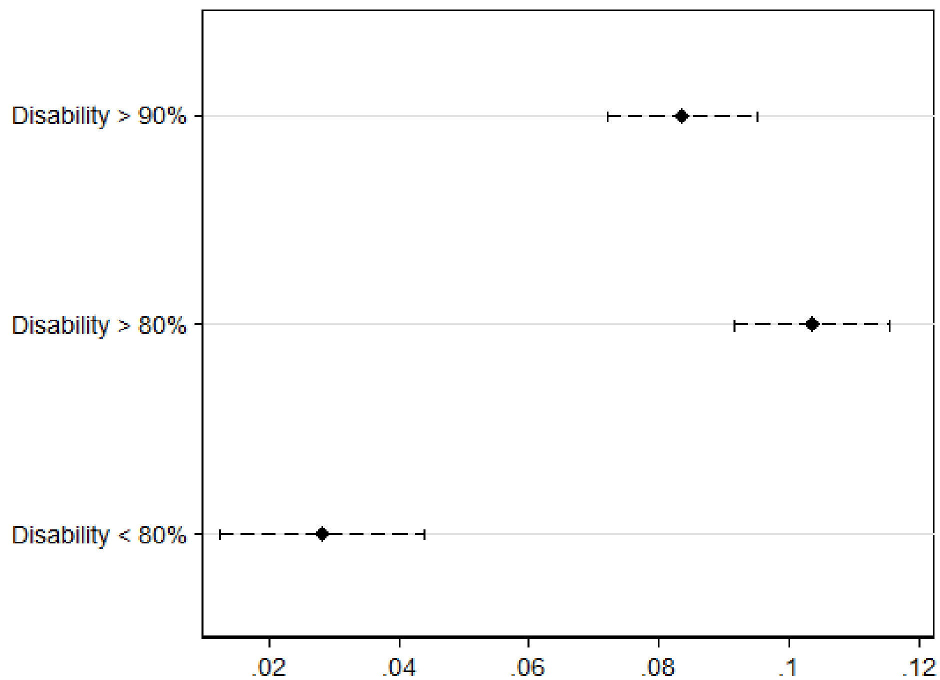
The graph shows the coefficients of the variable "Partner caregiver" discriminated by disability level, in the SAH regression. The model includes both individuals and partner characteristics.

Figure C.5: Partner caregiving and depression by disability



The graph shows the coefficients of the variable "Partner caregiver" discriminated by disability level, in the depression regression. The model includes both individuals and partner characteristics.

Figure C.6: Partner caregiving and disability by disability



The graph shows the coefficients of the variable “Partner caregiver” discriminated by disability level, in the disability regression. The model includes both individuals and partner characteristics.

Table C.3: Event study ($NC \to C$) - Caliper Robustness

	Self Assessed Health (OProbit)	Depression (OProbit)	Disability Index (GLM)
Caliper=0.05			
Treatment period	-0.005	0.040	0.022
	(0.033)	(0.032)	(0.014)
NC $\to$ C	0.036	0.038	0.048**
	(0.035)	(0.034)	(0.015)
NC $\to$ C in treatment period	0.078*	-0.015	0.012
	(0.039)	(0.039)	(0.018)
Caliper=0.1			
Treatment period	-0.005	0.040	0.022
	(0.033)	(0.032)	(0.014)
NC $\to$ C	0.036	0.037	0.048**
	(0.035)	(0.034)	(0.015)
NC $\to$ C in treatment period	0.080*	-0.015	0.013
	(0.039)	(0.039)	(0.018)
Caliper=0.2			
main			
Treatment period	-0.005	0.041	0.021
	(0.033)	(0.032)	(0.014)
NC $\to$ C	0.034	0.037	0.047**
	(0.035)	(0.034)	(0.015)
NC $\to$ C in treatment period	0.078*	-0.015	0.015
	(0.039)	(0.039)	(0.018)
Caliper=0.5			
Treatment period	-0.004	0.042	0.022
	(0.033)	(0.032)	(0.014)
NC $\to$ C	0.035	0.037	0.048**
	(0.035)	(0.033)	(0.015)
NC $\to$ C in treatment period	0.078*	-0.016	0.014
	(0.039)	(0.039)	(0.018)

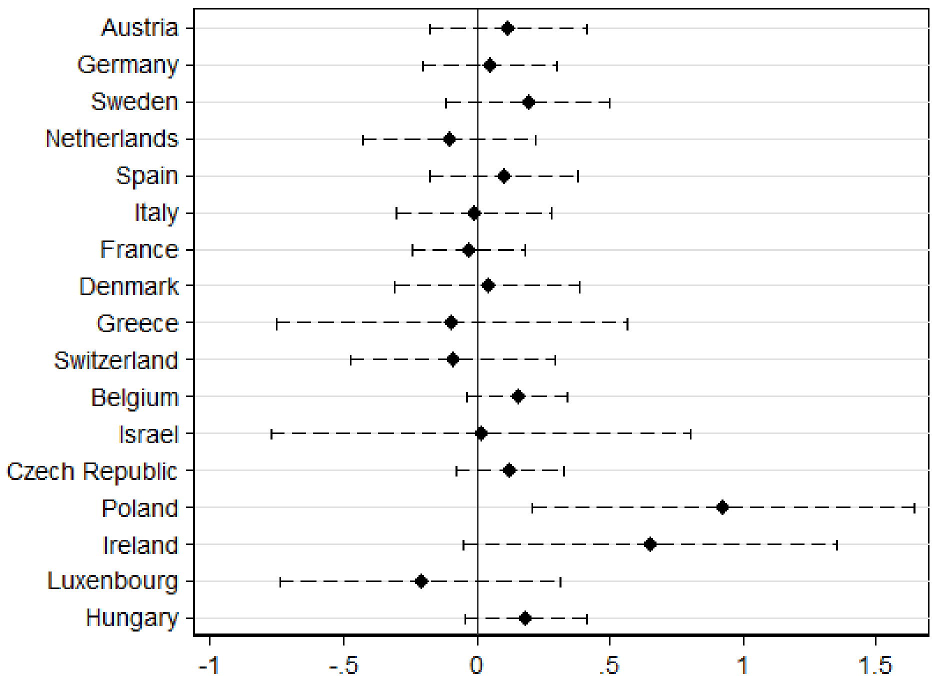
Regressions use a total of 7.704 individual surveys (from which 3,852 belong to the groups $NC \to C$) collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 17 countries. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. All individual and partner characteristics defined in Section 3.3.1 are used in the regressions but are omitted in the table.

Table C.4: Event study ($C \to C$) - Caliper Robustness

	Self Assessed Health (OProbit)	Depression (OProbit)	Disability Index (GLM)
Caliper=0.05			
Treatment period	0.002	0.015	0.010
	(0.075)	(0.076)	(0.032)
C → C	0.250***	0.103	0.083*
	(0.076)	(0.074)	(0.034)
C → C in treatment period	-0.076	-0.018	-0.021
	(0.077)	(0.078)	(0.033)
Caliper=0.1			
Treatment period	-0.007	0.012	0.012
	(0.075)	(0.076)	(0.032)
C → C	0.247**	0.112	0.082*
	(0.076)	(0.074)	(0.033)
C → C in treatment period	-0.080	-0.022	-0.021
	(0.077)	(0.078)	(0.033)
Caliper=0.2			
Treatment period	-0.012	0.012	0.015
	(0.074)	(0.075)	(0.031)
C → C	0.252***	0.127	0.094**
	(0.076)	(0.073)	(0.034)
C → C in treatment period	-0.084	-0.020	-0.027
	(0.076)	(0.077)	(0.033)
Caliper=0.5			
Treatment period	-0.021	0.016	0.015
	(0.074)	(0.074)	(0.031)
C → C	0.237**	0.128	0.095**
	(0.075)	(0.072)	(0.034)
C → C in treatment period	-0.047	-0.026	-0.029
	(0.076)	(0.076)	(0.032)

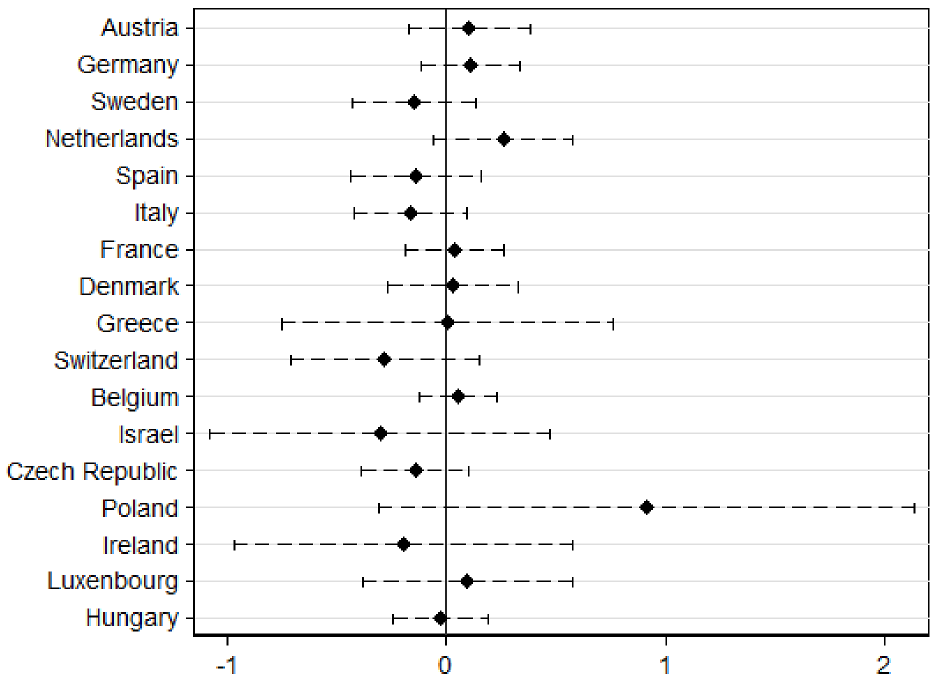
Regressions use a total of 1.848 individual surveys (from which 924 belong to the groups $C \to C$) collected in waves 1, 2, 4, 5 and 6 of SHARE (2004, 2007, 2011, 2013 and 2015), in a total of 17 countries. Standard Errors are clustered at the household level and shown in parenthesis. All models include fixed effects at the country and wave level. All individual and partner characteristics defined in Section 3.3.1 are used in the regressions but are omitted in the table.

Figure C.7: Effect of transition into caregiving on SAH by country



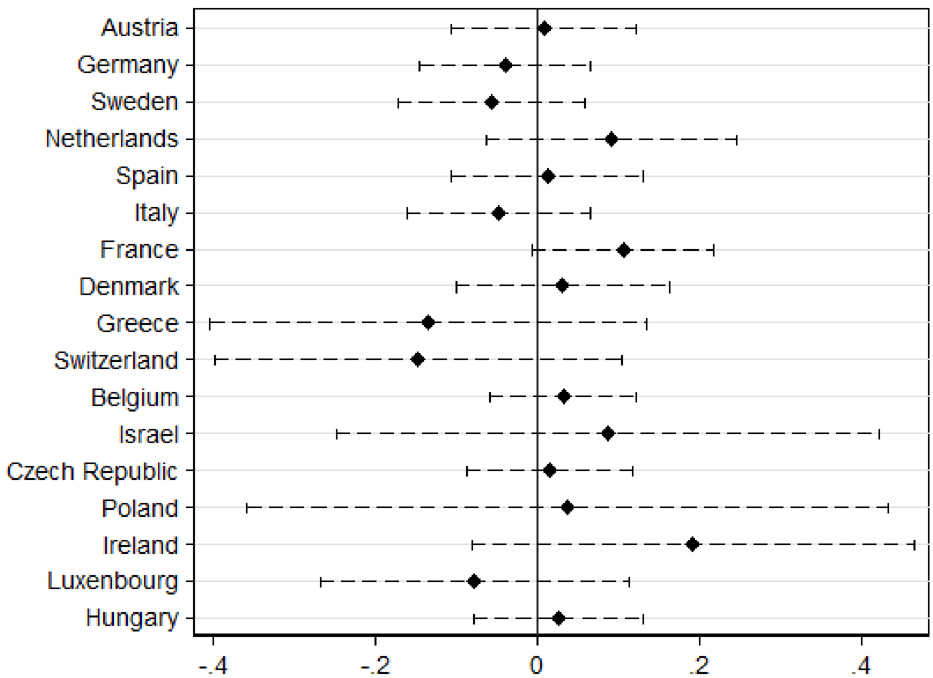
The graph shows the effect of transitioning into partner caregiving ($NC \to C$) discriminated by country, in the SAH regression.

Figure C.8: Effect of transition into caregiving on depression by country



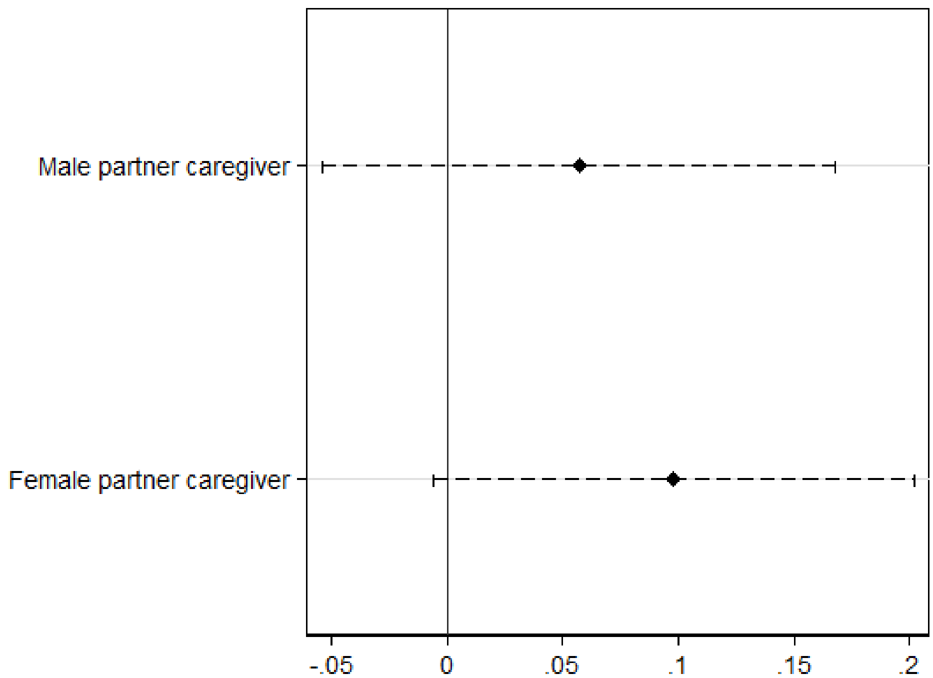
The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by country, in the depression regression.

Figure C.9: Effect of transition into caregiving on disability by country



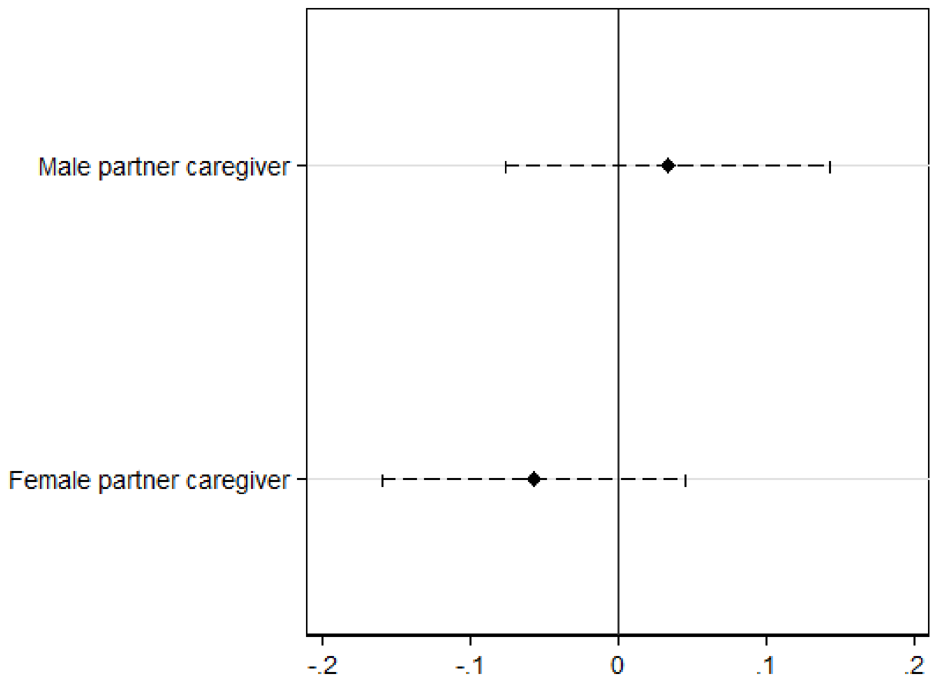
The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by country, in the disability regression.

Figure C.10: Effect of transition into caregiving on SAH by gender



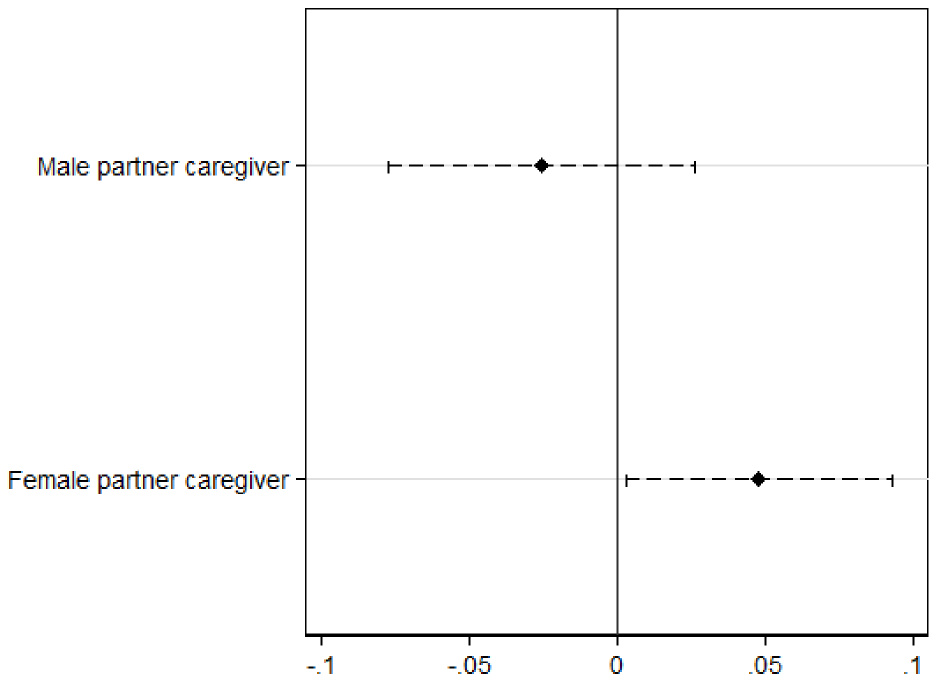
The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by gender, in the SAH regression.

Figure C.11: Effect of transition into caregiving on depression by gender



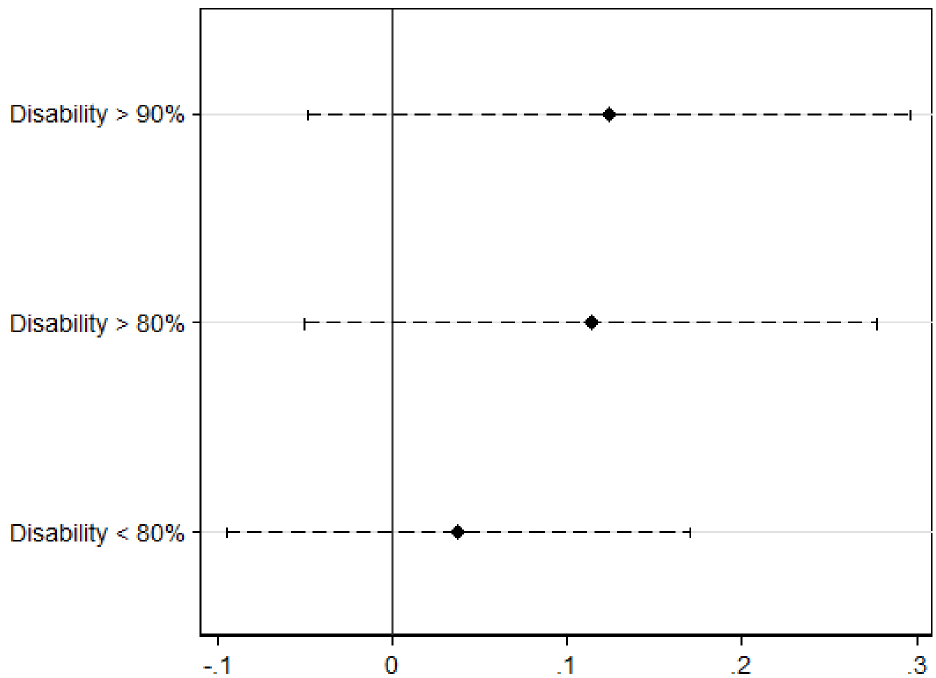
The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by gender, in the depression regression.

Figure C.12: Effect of transition into caregiving on disability by gender



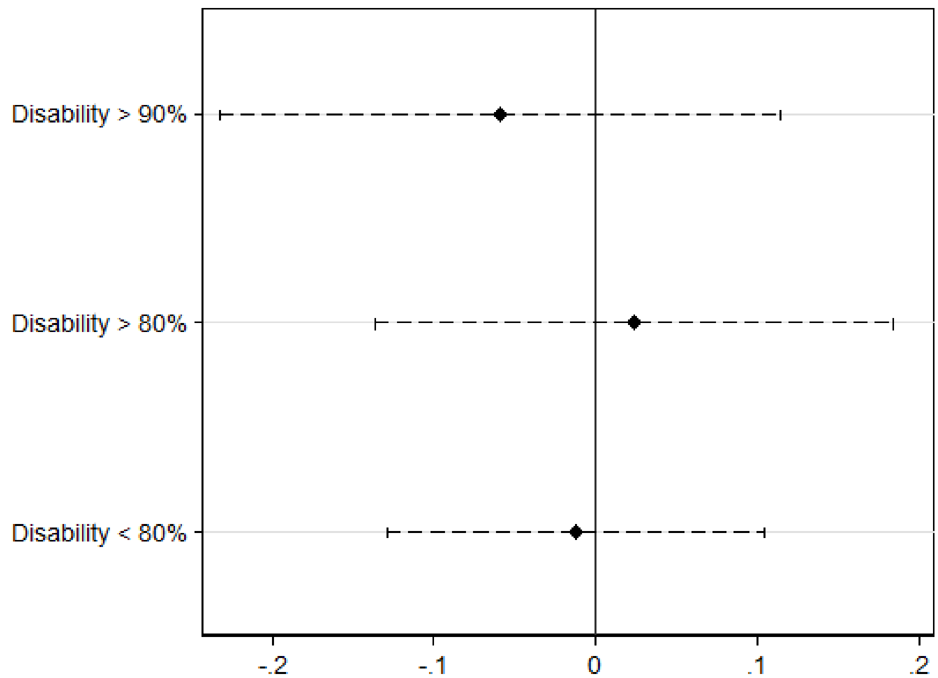
The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by gender, in the disability regression.

Figure C.13: Effect of transition into caregiving on SAH by disability



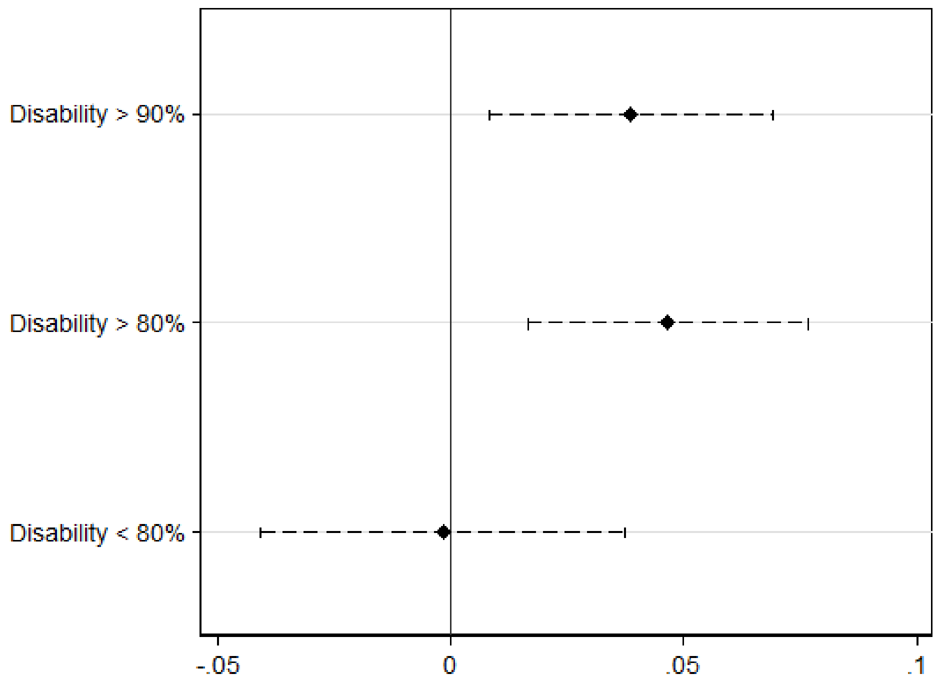
The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by disability, in the SAH regression.

Figure C.14: Effect of transition into caregiving on depression by disability



The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by disability, in the depression regression.

Figure C.15: Effect of transition into caregiving on disability by disability



The graph shows the effect of transitioning into partner caregiving ($NC \rightarrow C$) discriminated by disability, in the disability regression.