RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

IFN Working Paper No. 1365, 2020

Quality and Efficiency Between Public and Private Firms: Evidence From Ambulance Services

Daniel Knutsson and Björn Tyrefors

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Quality and Efficiency Between Public and Private Firms: Evidence From Ambulance Services

By DANIEL KNUTSSON AND BJÖRN TYREFORS *

THIS VERSION: 2020-09-04

Abstract

Economic theory predicts that outsourcing public services to private firms will reduce costs, but the effect on quality is ambiguous. We explore quality differences between publicly and privately owned ambulances in a setting where patients are as good as randomly assigned to ambulances of different ownership statuses. We find that privately owned ambulances are better at responding to contracted quality measures but perform worse on noncontracted measures, such as mortality. In fact, a randomly allocated patient has a significantly higher risk of death if a private ambulance is dispatched. We also present suggestive evidence on the mechanism, supporting that private firms cost innovate at the expense of ambulance staff quality.

Keywords: Public outsourcing, Pre-hospital care, Healthcare quality, Health

JEL classifications: P48, H44, I11, D22, D44, L33

* Corresponding author: Tyrefors: Research Institute of Industrial Economics (IFN), Box 55665, 102 15 Stockholm (e-mail, bjorn.tyrefors@ifn.se; telephone: +46(0)8-665 4500. and Dep. of Economics, Stockholm University; Knutsson: IFN, P.O. Box 55665, SE-10215 Stockholm, Sweden (e-mail, daniel.knutsson@ifn.se; telephone: +46(0)8-665 4500). We thank the Swedish Competition Authority and Jan Wallanders och Tom Hedelius stiftelse for financial support and Andreas Madestam, Per Pettersson-Lidbom, Ekaterina Zhuravskaya, Magnus Johannesson, Erik Lindqvist, Jonas Vlachos and seminar participants at IFN and Stockholm University for helpful comments. We are further indebted to Ola Kostenniemi (at SOS Alarm AB) and Gunnar Ljunggren (at SLL) for helping us access the data. Moreover, Ann-Catrin Eriksson has provided us with important information on the firms and contracts. Finally, Alexander Näsström assisted us in working with employer-employee data. We wish to thank all of you for your valuable assistance.

1 Introduction

Whether the private sector can provide public services more efficiently than the public sector is currently a central question in economics. Outsourcing to private firms is a common way for governments to supply goods and services to their citizens, and in many OECD countries, welfare services are increasingly being contracted out to private providers. Among OECD countries, on average, 3% of GDP is used for public procurement that is not consumed by the government itself, e.g., through outsourcing. Despite this strong reliance on outsourcing, theoretically, it is not a panacea for public sector inefficiencies.

When private firms have "residual control rights" over assets used to produce goods and services, they have incentives for cost saving innovation that the public sector does not (Hart and Moore, 1990). However, the canonical literature also points to circumstances where private firms have incentives to reduce costs by reducing quality (Hart et al., 1997). Cost innovation at the expense of quality is more likely to appear when quality cannot be adequately contracted on, innovation pressure is low, competition is weak, consumer choice is inefficient, and reputational mechanisms are weak (Schleifer, 1998). Moreover, another strand of literature suggests that many public sector activities are mission-oriented, where employees are intrinsically motivated, and it is therefore less costly to provide incentives in the public domain (Besley and Ghatak, 2005).

Thus, the advantage of outsourcing public services is ultimately an empirical question, but credible evidence is scant.[1] In regard to experimental evidence, there are only a few studies on the effects of privatization or outsourcing. The general results point to very

[1] There are many studies using non(quasi)-experimental methods in areas such as garbage collection (e.g., Szymanski and Wilkins,1993, Szymanski, 1996), prisons (e.g., Farabee and Knight, 2002, Duwe and Clark, 2013, Powers et al., 2017, Spivak and Sharp, 2008), and residential youth care (e.g., Bayer and Pozen 2005, Armstrong and MacKenzie, 2003).

small, if any, effects.[2] In particular, the few papers investigating healthcare markets reflect the inherent problems of self-selection and measurement quality. Most of the papers that are related to healthcare study quality differences between different nursing home providers in general and rarely outsourcing in particular.[3]

This paper provides, to the best of our knowledge, the first credible empirical evaluation of quality in performance between public and private firms in an acute healthcare environment. Analyzing the efficiency of emergency care is of economic importance, as aggregate spending on emergency care could be as high as 10% of national health expenditures (Lee et al. 2013). Interestingly, this is a setting where quality is inherently difficult to contract on. One important reason for contracting difficulties is because free choice of provider allows for self-selection of patients to providers. When individuals are allowed to self-select to private or public alternatives, separating the effect of ownership

[2] Bennmarker et al. (2013) find no general differences in the relative performance of private and public job placement agencies in Sweden. Behaghel et al. (2014) find in France that the public provider increased employment to a substantially larger extent, and in Laun and Skogman Thoursie's (2014) studies of the privatization of vocational rehabilitation, they find that there are no large efficiency gains from privatizing vocational rehabilitation. Mukherjee (2020) find that inmates in private prisons spent 90 days longer in prison, eroding the cost saving by half. Jansson and Castwall (2017) used part of one year of our data as a pilot for a bachelor thesis under the supervision of Tyrefors.

[3] Research on nursing homes, where quality is measured as health outcomes, show mixed results. Private firms have worse outcomes in some settings (Barron and West, 2017; Tanuseputro et al., 2015) and better outcomes in other contexts (Bergman et al., 2016). Other studies have found that private firms give priority to providing service to the residents of nursing homes, but at the same time have a lower staff to patient ratio (Stolt et al., 2011). Comparing for-profit with nonprofit nursing homes, Grabowski et al. (2013) found that the for-profit nursing homes had worse health outcomes. A similar result has been found in the hospital market (Devereaux et al., 2002). However, in more recent work, Wubker and Wuckel (2019) found that private for-profit hospitals in Germany performed better than other hospitals for pneumonia patients but no difference was found for heart attack patients. Duggan, Gruber, and Wabson (2018) study private health insurance plans competing with Medicare and find that healthcare utilization increased when the private alternative exited the market, without leading to any discernible effect on health.

from patient-specific characteristics is challenging. Furthermore, quality is often difficult to measure due to small samples. Thus, available measures may be imperfect, unverifiable, or observable only to the provider.

Methodologically, we use the ambulance services in Stockholm (Sweden) as a laboratory to answer questions of outsourcing in healthcare. The primary methodological challenge we face is that both ambulance providers and patients are geographically clustered. Our empirical approach to handle this obstacle is similar to but distinct from that of Doyle et al. (2015, 2017). We compare patients living in the same residential block and where ambulance assignment between private and public firms is determined in a stochastic way by ambulance availability and ambulance distance to the patient at the time of the call.

Importantly, the provider type is chosen independently of the patient characteristics. Thus, there are no confounders, and we can credibly identify causal effects of provider type on quality. This is verified by a test of balance and correspondence about criteria with the assignment center. There is no evidence of a systematic relation between the assignment of the type of provider and patient characteristics.

Our first set of results relates to firm-specific ambulance outcomes of each patient serviced. These outcomes are all explicitly contracted upon and well documented, ultimately with fines if inferior performance is observed. In terms of contracted, salient outcomes, we find that private firms are more efficient. Private firms were 8% faster in responding to a dispatch and 7% faster in reaching the patient. These results indicate that private firms responded to the incentives provided by contracts and the follow-up meetings with explicit feedback. We interpret these findings as private firms being more innovative than public firms when outcomes are observable, measurable, and contracted upon.

In our second set of results, we find that private ambulances produced *lower* quality with respect to noncontracted outcomes (i.e., mortality). Privately serviced patients had a

1.4% higher chance of dying within three years of service. This mortality penalty for private firms is consistently higher *from the first day* of service and up to three years later. A back of the envelope calculation for a given year shows that approximately 400 additional patients, in Stockholm alone, died within 3 years due to the ownership status of the ambulance transporting them. This amount outnumbers the traffic deaths in all of Sweden in a single year by more than 50%. Considering the two sets of results together, we argue that our findings provide strong support for the theoretical predictions in Hart and Moore (1990), Shleifer (1998), and Hart et al. (1997).

We argue that private firms have a different recruitment strategy and that staff quality is one likely reason for the mortality differences. This hypothesis is supported by data. When they reach the patient, the ambulance crew members make their own judgment of the injury. Private firm ambulances assessed their patients as being less severely injured, although injury status was *ex ante* balanced across providers. This finding is interesting given that private firm patients have higher mortality in both the short and long terms. Consistent with this assessment, private firm ambulances left 25% more patients at home (3 ppt out of an average of 12%), which could explain some of the increased mortality risk already at day one. These differences between firms are corroborated by descriptive evidence. In assessments of each firm from 2017, there are indications that private firms have, on average, both a higher turnover rate and a higher fraction of their staff on temporary employment contracts (SLL, 2017). Additional information from matched employer-employee data at the regional level suggests that private firms have a different recruitment strategy as employees are younger, have less experience, change workplace more often, and have lower grades. Moreover, anecdotal evidence lines up well with this description.

Our calculations suggest that the cost of private firms in a given year in terms of loss of quality adjusted years (QALY) could be substantial. In relation to the *total* cost of the ground-based ambulance services in Stockholm, a conservative estimate puts the annual loss of QALY *at 50%* of the total annual cost.

Our findings contribute to the literature on outsourcing and quality, surveyed recently in Andersson et al. (2019), and support that the ability to contract quality is paramount to the success of outsourcing public services. Our paper adds to the small literature stream on quality outcomes from private firms producing healthcare for the public sector (Bergman et al., 2016). We confirm several predictions of the theory presented by Hart et al. (1997) in a setting with the lowest bid auctions as the allocation mechanism. First, uncontracted (unmeasured) quality tends to deteriorate as private firms cost innovate. Second, private firms have incentives to uphold quality that is contracted. These findings provide policy-relevant information on when and how outsourcing can be successful, in terms of both cost and quality.

The findings of this paper also add to the literature on the importance of healthcare resources and regulation on patient outcomes (see, e.g., Athey and Stern, 2002; Doyle, 2011; Doyle et al., 2015, 2017).[4] Especially relevant are the few papers studying staff productivity, which can be crucially important in healthcare. Bartel et al. (2014) found that

[4] The ambulance sector has received much attention in the last few years as a setting where a treatment assignment mechanism can be observed. Doyle et al. (2015) use the fact that different ambulance firms have different preferences regarding the hospital to take patients to. Together with ambulances being as-good-as randomly assigned to each patient (through competition of rotating dispatches), the authors can estimate the effects of hospital quality on patient outcomes. They find that hospitals that receive higher payments from Medicare have better outcomes. Other similar papers have additionally used the same empirical framework and setting to inform policy makers about, e.g., how to measure hospital quality (see, e.g., Doyle et al., 2017; Hull, 2018).

nurses' education and on the unit experience increased productivity in hospital environments. Our findings suggest that even important health outcomes can be susceptible to healthcare staff quality. Moreover, Gruber and Kleiner (2012) found that during nurses' strikes in New York, in-hospital mortality increased by 18%, even when hospitals engaged temporary staff. Similarly, Gruber et al. (2018) study physician practices in emergency departments (EDs) as a new regulation imposed a strict ceiling on waiting time before seeing a physician. As patients were taken care of earlier, more patients were admitted to a hospital ward, and mortality decreased by 14%. Our findings confirm the susceptibility of the group of patients seeking care at the ED to small changes in staff practices on mortality.

The rest of the paper is organized as follows. The next sections address the background and institutional details, followed by section three, where we describe the methodology and empirical design. In section four, the results are presented, and in section five we provide a cost figure of the inefficiency. In section six, we conclude.

2 Background

2.1 The Ambulance Market in Stockholm

The Stockholm region is the largest healthcare region in Sweden, with 2.3 million inhabitants in 2017. In 2016, the inhabitants in the region of Stockholm were serviced by 73 ambulances that responded to 220,000 dispatches at a total cost of 520 million SEK (approximately 52 million Euro). From 2010, the number of ambulances increased over time through supplementary contracts reflecting the increasing population of Stockholm. The total costs of ambulance ground transports were approximately 360 million SEK in 2010, when only 55 ambulances took on 145,000 assignments (Hälso- och sjukvårdsförvaltningen, 2011, 2018).

From being completely run by a public firm, in 1993, ambulance services were introduced to competition. The service has ever since been financed by the public sector (through regional income taxes) but with firms competing for contracts in lowest bid auctions spanning approximately six years.[5] The county of Stockholm was initially divided into 7 districts that the firms competed for. Several firms have entered and disappeared over time, and in 2006, it was decided that the public firm was to be freed from competition and provided with two districts not up for auction, located in the central part of Stockholm city. Private firms have since competed for five districts, with each firm allowed to operate at most three out of the five (AISAB, 2020). In 2009, there were three private firms operating five districts of Stockholm in addition to the public firm operating its two districts. During the time period that we study, only one auction took place (in 2011). During this auction, one of the firms lost its contract, meaning that only two private firms remained. Currently, the two remaining firms still comprise the market structure even though the auctions offer (conditionally) free entry and other firms have tried to compete. Figure 1 shows the locations of the stations during this time period.

Figure 1: Location of Ambulance Stations in Stockholm County (about here)

In auctions before 2011, winning the right to operate a sector was a function of being able to comply with the rather precisely specified contract at the lowest cost. The contract specified a fine if the response times after dispatch calls were too slow. Ambulance crews were to be ready to depart within 90 seconds after a dispatch for a high-priority case

[5] The public ambulance organization was reorganized in 1993 as a publicly owned incorporated firm that competed for contracts on equal terms to the private firms. Contracts last for five years, with the additional possibility of an extension up to two years. Auctions for ground-based ambulances were held in 2005, 2011, and 2017.

(priority 1) and within 180 seconds for an intermediate priority case (priority 2 or 3). Fines could also be generally applied if the contracted ambulances were not operationally sufficient (on repair without an extra car or without staff) or if staffing requirements were not fulfilled. Time to patient was carefully monitored but not tied to a monetary fine. Ambulance firms were compensated with a fixed and variable component. The fixed amount, allowing firms to keep ambulances, staff, and a garage, was adjusted on a yearly basis for changes in the number of transports compared to the previous year. More assignments led to marginally higher compensation (AISAB, 2005).

From 2011, the county of Stockholm started to weight quality as 25% of the bid, but otherwise, the same structure was maintained with respect to fines. Another difference was that the variable compensation was removed, and requirements to gradually increase the use of environmentally friendly fuels were tied to fines. The maximum fines that firms could be required to pay were 3% of the contract value (Landstingsrevisorerna, 2013). Fines were not conditioned on health outcomes or mortality. The contracts also specified the number of ambulances in each area, their equipment, and their service time availability. Additionally, from 2008, a registered nurse with a master's degree in nursing was required to attend an emergency ambulance together with at least a trained paramedic.[6] When

[6] In addition to including the general staff requirement of regular ambulances, the contracts included other types of units. From 2009 to 2012, these other units consisted of several transport ambulances that were mainly staffed by paramedics and responded to low-priority dispatches without any evaluated medical treatments required. There were also two emergency cars in Stockholm. These vehicles included a nurse with a specialist degree in anesthesiology and were not able to transport patients, only to assist an ambulance with critically ill patients who might require intubation. Furthermore, two helicopters were contracted, of which one was available during the summer months (Hälso- och sjukvårdsförvaltningen, 2011). Under the new contract from 2012, the ambulance system changed somewhat. An emergency car was instituted with a specialist physician onboard. The two emergency cars that were already operational were updated to patient transportation capacity. This meant that these units could take on emergencies by themselves if required and transport patients to hospitals. The staffing of these ambulances was unchanged, and they were still intended to be used in medical emergencies and in support of other ambulance units. From 2012, the helicopters were

studying the continuous evaluations between the Region of Stockholm and the ambulance providers, these factors are mentioned, measured, and discussed.

2.2 The Geography of Ambulance Transports in Stockholm and the Dispatch

The ambulance dispatch firm in Stockholm, SOS Alarm AB, has two main objectives. First, patients should be prioritized such that they can be served based on the severity of their condition. Second, patients are matched to ambulances with as little waiting time as possible before the ambulance arrives while still maintaining sufficient readiness for other medical emergencies. This objective is most crucial for critically ill patients with, e.g., an ongoing heart attack or recent stroke and for whom time to treatment is of the essence (SOS, 2020. Webpage).[7]

When calling the emergency number 112, the caller is connected to an emergency operator. Based on the sort of emergency, the caller is further connected to a specialist (normally not a health professional) to further assess the emergency. Once the specialist has made an assessment of the patients' needs (severity), an ambulance work description is sent out, including the severity of the emergency, patient characteristics, characteristics of the emergency, and patient location. A dispatch operator then takes the job and assigns the patient to an ambulance according to the main objectives. In most cases, the dispatch operator observes the patients' location and priority (on a scale from 1-4 where priority 1 is the highest) when assigning an ambulance to the patient. According to SOS Alarm AB, the assignment should immediately be directed to the closest ambulance if the patient's condition is life threatening (priority 1) or potentially life threatening (priority 2). The assignment can be put on hold if no suitable units are available (priority 3). In reality, cases

supplied by a separate private firm, not operating ground transports, while the emergency ambulance cars were supplied by the private firms (Hälso- och sjukvårdsförvaltningen, 2018).

[7] Information regarding SOS Alarm AB and how it works has primarily been collected through personal contacts with employees in the firm.

with priority 2 or 3 can be put on hold to facilitate readiness for emergencies in the entire region (Riksrevisionen, 2014. pp. 36).

Within each firm-specific district, there can be several ambulance stations. These stations, containing one or more ambulances, are connected to a geographical area in which they are primarily responsible for transporting patients. If the local ambulance is unavailable in a district, e.g., occupied by another patient or relatively far away from the patient, another ambulance is called in to transport the patient based on proximity. Dispatch services are explicitly not limited by the boundaries formed around each ambulance station, and dispatch operators can freely engage ambulances as they find appropriate. An ambulance can be dispatched to any location in the county, although very long transport times to patients are rare.

In essence, whether a patient is picked up by the local ambulance or by one from a nearby station depends on the relative distance and availability of that ambulance compared to others nearby. Low priority cases, whose condition has been evaluated not to be affected by waiting, can be put on queue until the local ambulance is available again or assigned to another ambulance nearby. This will depend on the general demand for ambulance transport (readiness for emergencies) and the patients' waiting time in the queue. We use these features of the ambulance allocation mechanism to design a natural experiment to evaluate the causal effects of the ambulance provider.

Importantly, two patients residing in the same neighborhood have the same assigned hospital. The ambulance provider takes them to their designated hospital, determined by their residence, or to another hospital where the patient is currently treated or where specialist care that the patient requires is available.[8] However, this is not the choice of the

---

[8] In Stockholm, residents are treated at hospitals based on where they live. This is intended to facilitate a reasonably predictable demand at each emergency department. Since 2016, there has been discussion around

ambulance provider. Therefore, distance to hospital or hospital resources are not mechanisms that can explain any results that we find, in contrast to, e.g., Doyle et al. (2017).

3 Empirical Design and Data

3.1 Empirical Design

Estimating the causal effects of outsourcing public health services to private firms on quality outcomes is inherently challenging. The main challenge comes from patients being able to sort themselves to different health providers based on social and material resources, such as income, education, personal contacts, and residence. Furthermore, patients' health may determine by whom and at what location they receive care.

The ambulance services that we study provide some clear benefits with respect to this type of sorting. Patients cannot sort themselves to either a private or a public ambulance. The dispatch operator assigns with certainty the patient to an ambulance based on criteria that the patients themselves cannot affect. Dispatch operators are responsible for minimizing the waiting time for patients while assuring that the time to a future patient in distress is minimized (readiness for emergencies). This is done manually by observing a digital map with ambulance resources and patients marked. For severe cases (priorities 1 and 2), distance (or time to patient) is the sole ambulance allocation mechanism. Furthermore, ambulances are, in general, not different in terms of resources. All ambulances have a fixed set of equipment and rigid requirements of formal staff quality.[9]

free choice of provider in terms of emergency departments. Nevertheless, as of 2020, this has not been fully adopted in the ambulance services.

[9] Before 2012, not all ambulances were equipped with ECG monitoring capabilities. Only ambulances operating around the clock had this equipment. These ambulances were not exclusive to any particular provider. Formal staff quality means that there were requirements on the education and certification of the staff operating ambulances. Before 2012 transport ambulances had a lower staff requirement, at least 2

However, even if patients cannot choose the ambulance providers, they choose where to live, and providers operate more frequently in certain districts (primary uptake areas). The public firm mainly services the inner city and a few nearby areas. The private firms mainly service the suburbs and distant cities, villages, and rural communities in the larger Stockholm region. These populations are likely very different in terms of many characteristics. A naive comparison of outcomes by public-private status is, hence, likely to yield a biased estimate of efficiency due to noncomparable treatment and control groups. Fortunately, the ambulance assignment is not bound by the district borders. This means that almost all patients have a nonzero probability of being serviced by a private (public) ambulance regardless of residency.

To identify causal effects of ownership status of ambulances in this setting, we rely on the cross-district movements of ambulances.[10] There are several different situations that give rise to cross-border movements. First, the district-specific ambulance can already be servicing a patient and thus be unavailable for a new assignment. Second, an ambulance from a different district might be closest to the patient and thereby the best choice given that the patient is in need of urgent assistance. Third, maintaining readiness for coming emergencies might require assigning a different ambulance than the district-specific ambulance. These situations are unpredictable and unrelated to patient health given the priority set by the emergency operator and the location of the patient.

To be able to use movements of ambulances across districts, we take a nonparametric approach by controlling for neighborhood fixed effects. By using detailed information on each patient's location with GPS coordinates, we construct a grid of squares covering the

paramedics, which changed after 2012 to one paramedic and one specialist nurse. From 2009 transport ambulances with lower staff requirements were phased out.

[10] Our design also leverages the variation across district borders, but these often coincide with natural barriers, such as water bodies, highways, or sparsely populated areas.

full county of Stockholm. Patients requesting an ambulance from a given grid cell are assigned to a group variable that we use as fixed effects. With these fixed effects, we are essentially comparing patients who live very close to each other.[11] Our baseline grid cells are approximately 800 x 800 meters, resulting in around 7,000 geographical fixed effects with variation in the ambulance provider.[12]

These patients, by virtue of their residence, have a very similar probability of being serviced by a private (public) ambulance. By holding the probability of a private ambulance constant, we use the quasi-random variation in ambulance assignment generated by idiosyncratic shocks to local ambulance demand and ambulance locations at the time of the call. Put differently, the theoretical experiment that we propose is the following. Several patients live close to each other and call for an ambulance at different times for different conditions. Some of these patients are serviced by a public ambulance due to unpredictable circumstances, such as those described above, even though they reside in a private ambulance district, while others are serviced by a private firm.[13] When we reduce the size of the geographical unit in which we compare patients, the probability of a patient being

[11] Our baseline specification compares patients living up to 800 meters from each other. In our most conservative specification, we compare patients living approximately 100 meters from each other, and the main results are robust. As we have access to coordinates of patients at the time of the emergency call, these do not perfectly represent residence but can also describe the workplace or other accident locations.

[12] The 7,000 fixed effects described in the text are the effective number used in estimation. This means that this is the number of geographical areas where we observed patients serviced by both public and private ambulances in the sample period.

[13] Essentially, this probability depends on which ambulance district the neighborhood is situated in (private/public) and its location in relation to other districts (private/public). A private district residential block bordering a public firm district has a higher probability of being serviced by a public ambulance than a residential block close to another private district and far from any public ambulance districts.

assigned a private ambulance essentially converges to a constant, and the assignment of ambulances, in terms of ownership status, is not related to patient health.[14]

This design further provides benefits for identification. Comparing patients living close to each other means that the patients are not only faced with a similar probability of treatment but also that they should be transported to the same hospital from approximately the same transport distance.[15] Therefore, distance to hospital or hospital resources are not mechanisms that can explain any results that we find. Moreover, patients residing close to each other are similar to each other in a range of dimensions, allowing us to estimate precise coefficients, as our fixed effects can explain much of the residual variation.

In addition to geographical fixed effects, we include time dummies (year and month dummies) in our main specifications, and we include dummies for how urgent the dispatch operator determines the assignment to be.[16] Transport urgency is part of the observable treatment allocation mechanism, and it may be important to hold it constant. Similarly, with time effects, we can hold constant changes in the treatment probabilities over time that result from introducing new ambulances to different providers. Including these control variables, our main model to be estimated by OLS is

---

[14] Importantly, ownership status of each ambulance has no impact on the dispatch choice, and firms receive no marginal monetary gain for transporting a patient. We also show that patients serviced by either private or public ambulances, but residing in the same neighborhood, are indistinguishable from each other (balanced) in terms of age, sex, prior healthcare visits, and diagnoses, such as diabetes, hypertension, and cardiovascular disease.

[15] Some patients do not live where they are picked up for transport to hospital, and there are a few exceptions. However, the provider does not decide on the destination.

[16] Transports are defined in four categories determined by the urgency of the transport, as follows: 1 - Immediate threat to life, 2 - Urgent, 3 - Transport with medical requirements but where time is not essential to patient health, and 4 - Patient who has to be transported lying down but who does not require any medical assistance diagnoses recorded during these visits.

$$Outcome_{igy} = \alpha_g + \beta * Private_{igy} + Year_y + Month_y + Priority_{igy} + \epsilon_{igy} \qquad (1)$$

In the above equation, the outcome of individual i in grid cell g at time y is causally related to the ownership status of the ambulance servicing the patient. Clearly, the aim of this natural experiment is to mimic a random allocation of patients to ambulances. Assuming that the treatment is as good as randomly allocated within the grid cells, with a different, but within-group constant, probability of treatment, we have to control for fixed effects at the group level (see, e.g., Krueger, 1999). A testable implication of our design is that we expect to see similar characteristics between patients in the same grid cell serviced by either private or public firms.

To test whether patients serviced by private or public firms are similar before the dispatch, we use predetermined variables on the left-hand side in equation (1) and estimate the association with our *private* dummy variable (the treatment). The estimated β is now the difference in these characteristics between patients serviced by a public and a private ambulance conditional on the control variables included. The predetermined variables that we consider are measured either at the time of ambulance dispatch (age, sex, previous ambulance transport, and residence in a nursing home) or before the study period begins in 2007-2008. To further show that our design solves inherent imbalances between patients serviced by private and public ambulances, we also display correlations between the predetermined variables and the *Private* dummy, i.e. the results from a bivariate regression.

We further take a design-based approach to inference. In a randomized controlled trial where individuals are randomized, which is our target design, there is no general need for clustering (Abadie et al., 2017). Since the treatment is quasi-randomized at the individual level, with repeated independent experiments over time, we cluster our standard errors at the individual level. In our robustness analysis, we explore more conservative levels of clustering.

3.2 Data and Variables

In addition to offering a new and credible way of identifying differences between public and private firms for relevant outcomes, we contribute in this paper by observing and evaluating contracted and noncontracted quality indicators. Contracted indicators are time stamps in the data describing the time interval of the ambulance dispatch during all stages. Slow firm average response times, or slow times to patient, are under contract and associated with fines or feedback in follow-up meetings. Health outcomes are noncontracted and possibly noncontractible since contracting, for example, mortality requires large amounts of data (unless effect sizes are very large) and contractible measures that are robust to provider moral hazard (Gupta, 2017). In our data, we observe more than one million ambulance dispatches, where close to 60% are provided by private firms. These data include all inpatient and outpatient visits, mortality, and, importantly, information determined prior to the ambulance dispatch.

The data come from the VAL database, which originates from Region Stockholm. Region Stockholm is responsible for all publicly financed health care in greater Stockholm and keeps health registries for research, evaluations, and monitoring of the care that it finances. We make use of the ambulance registry and pool all ambulance dispatches between 2009 and 2016.[17] Then, we merge several other registries from the same database.

[17] For 2016, the data are incomplete due to the phasing in of a new ambulance report system called FRAPP. There is likely some additionally missing data in VAL regarding the ambulances, especially for the year 2011. However, the results are not sensitive to excluding this year. For more information on the VAL database, see http://www.gups.sll.se/val/default.htm

Inpatient care, outpatient visits, and mortality come from the same source. Additionally, we are able to merge patient GPS position data from the dispatch services, SOS Alarm AB, which we can uniquely merge to each ambulance assignment ID number.

In the data, we observe more than 1.1 million ambulance assignments for close to 500,000 unique patients. An ambulance assignment, i.e., an observation, is only registered in the database if a patient was present on the scene. We have information on time stamps for each dispatch, e.g., response time to a dispatch, time to patient, time at patient, time to hospital, and time at hospital.[18] Since some of these are important quality indicators and contracted on, they are used as outcomes to investigate how public and private firms behave differently during the assignment and respond to contracted incentives (where response time and time to patient are the most important indicators).

We make some data restrictions in the analysis by focusing attention only on emergency ambulances which are generally comparable in terms of staff and equipment requirements. This means that we are excluding transport ambulances (which were phased out from 2009), Intensive care ambulances (only publicly owned), and a few other rare ambulance types that are directed to particular patients (including Helicopters, that are strictly privately owned). Since the majority of units during this time are regular emergency ambulances, this exclusion amounts to a loss of 41,621 observations. We further exclude observations without a proper provider (1,000 observations) and extra ambulances that are used at times of especially high (but predictable) demand, such as midsummer and New Year's Eve (2,700 observations). Our final sample that we use consists of 1,099,980 observations at most.

[18] These variables had many missing observations in the original VAL database on some of the time stamps. This could be due to transfer problems from ambulances to the data system CAK-net, where the information was stored. We complement some time stamps with data directly from the dispatch services SOS Alarm AB, which keeps more complete data with respect to these variables.

The geographical position information we have for each patient is used to generate the grid cell fixed effects. To do this, we first transform the GPS data (latitude/longitude) to cartesian coordinates using the STATA command *geo2xy*. We then take out the two positions that describe and include the entire region of Stockholm (northwest and southeast) and generate the grid as X times X equally sized rectangles between these points by grouping observations in each cell (since Stockholm is not defined as a square, the sides are not exactly the same size). We do this for several different values of X, such as 200 and 1000, leading to rectangles with sides roughly between 800 and 100 meters.

As our main dependent variable on health, we use mortality. Since we have the date of death for all ambulance patients, we create several variables indicating whether the patient died within a certain length of time after each ambulance dispatch. The time periods that we consider are one day, one week, one month, three months, six months, one year, two years, and finally three years after each dispatch. Three years after an ambulance dispatch is somewhat of a stretch since our mortality data only reach up until 2017. Since the ambulance data stopped being recorded in the second half of 2016, three-year mortality is not properly coded (under reported) for patients serviced in 2015 and 2016. This measurement error is unlikely to be systematically related to firm ownership status and, hence, results only in inflated standard errors.

Table 1 shows descriptive statistics for the predetermined variables of the data used in the analysis. Forty-five percent of the patients had a previous ambulance contact in the data within one year, suggesting that the average serviced patient is prone to acute illness or has severe chronic health conditions. This can be explained by the ambulance serviced population being old, with the average patient being 62 years old and 7% of patients residing in a nursing home. They were also sick; 10% had a diabetes diagnosis, and 7% had a COPD diagnosis in 2007-2008. This picture is further reinforced in Table 2, where the outcomes are described. Ambulanced serviced patients have almost 30% three-year mortality rate and 10% one-month mortality rate. Although these numbers may be somewhat overstated because ambulance contacts could be more frequent at the end of life, it is clear that ambulance-serviced patients are often old, sick, and frail. Minor changes in their care or treatment could thereby have vast implications for their health.

Summary tables 1 and 2 (about here)

4 Results and Discussion

4.1 Balance

As noted in Figure 1, the ambulance stations are not randomly allocated, and an average person serviced by a private ambulance could be different from an average person serviced by a public ambulance. However, we argue that any imbalance across groups will reflect geographical differences in the probability of being serviced by a private (public) ambulance that are correlated with differences in patients' characteristics. If this is the case, any observed differences should vanish if conditioned on fixed effects, based on small enough geographical units (grids).

To emphasize this argument, Figure 2 and Figure 3 show the results from balance tests. In these figures, we use 25 predetermined characteristics as outcomes, including many health variables. In the left panel of both figures, we show the results where no grid fixed effects are included, i.e., the results from simple bivariate regression. In the right panel B, we use our designed based approach, the full specification of equation (1).

Starting with the left panel of Figure 2, we conclude that the patients where a private ambulance was dispatched are indeed unconditionally different. Private ambulance-serviced patients are, for example, 0.06 standard deviations younger than public ambulance- serviced patients. This amounts to a raw average age difference of 1.5 years. Reflecting this age difference, private ambulance-serviced patients seem, on average, to be healthier than public ambulance-serviced patients. They are less likely to live at a nursing home and had fewer health-care contacts in 2007-2008. The left panel of Figure 3 shows a more complicated picture, with private ambulance-serviced patients having more documented lifestyle-related diagnoses (such as diabetes, hypertension, and hyperlipidemia) but fewer diagnoses related to substance abuse.

In the right panel of Figures 2 and 3, we re-estimate the same models using our preferred specification (adding grid cell fixed effects, time dummies, and priority to patient). Out of 30 tested hypotheses only one, regarding whether the patient had any primary care contacts in 2007-2008, is significant at the 5% level. This is similar to what we would expect by random sampling. This estimate is negative and suggests that if anything, patients serviced by private firms were healthier prior to the dispatch. The other variables tested are close to zero and insignificant, providing strength to the assumptions underlying our design, i.e., ambulance firm assignment is unrelated to patient health conditional on the fixed effects. Furthermore, our preferred model explains much of the residual variation in the outcomes, as reflected by meaningful reductions in standard errors on all estimates. We interpret these findings as strong support for the ability of our design to handle the geographical selection as present in the data.

Figures 2 and 3. Balance of predetermined variables (about here)

As age could be the potential driver of many of the unconditional health differences, we provide kernel density plots of patient age distributions between public and private ambulance-serviced patients (see, figure 4). In the left panel of figure 4, we show the unconditional difference in age. We find that the distribution for public ambulance-serviced patients is skewed to the right, most likely because older people tend to live more centrally in apartments. Moreover, in the right panel, we residualize age by our preferred model (without the Private indicator) and estimate the same densities conditionally. We find that the two estimated distributions are conditionally extremely close to each other, further supporting our assumption of local randomization of patients to providers.

Figure 4 (about here)

We conclude that our empirical strategy seems plausible and delivers balanced treatment and control groups with respect to health and other predetermined outcomes.

4.2 Effects on Contracted Outcomes

We have information on two of the outcomes that are perfectly observable and contracted on: the response time after the dispatch call and ambulance time to patient. Again, we present the full specification of equation (1). In Table 2, column 1, we show that private ambulances respond 8 seconds faster, a decrease in response time of close to 8%. The standard errors in brackets show that the estimate is highly statistically significant, with a t-statistic of more than 10. This outcome is contracted and entails fines if performance is sufficiently poor. We further find, as reported in column 2, that private firms reach their patients faster. Private firms are 63 seconds (or approximately 8%) faster, which a considerable amount of time is given that time is a crucial success factor in the present setting. Again, the estimate is highly statistically significant. This outcome is not contracted with fines as the ultimate consequence but is an important indicator that is followed over time and discussed during follow-up meetings. Thus, the results confirm the predictions in (Hart and Moore, 1990 and Hart et al., 1997). Moreover, time delay is a key detrimental factor in the emergency literature (Jena et al. 2017; Lucchese, 2020).

Table 3. Contracted outcomes (about here)

4.3 Effects on Noncontracted Outcomes: Mortality

The ultimate quality measure for health services is often mortality. Again, we present the full specification of equation (1) now using mortality as the outcome. In Table 4, we present different mortality measures starting with 1-day mortality all the way to 3-year mortality. Column 1 starts displaying the effect of a private ambulance being dispatched instead of a public ambulance on the 1-day mortality. After one day, mortality among private ambulance-serviced patients is significantly higher (0.1 ppt). The effect increases rather monotonically when evaluating mortality over longer periods. That the effect is increasing over time, points to worse health outcomes both for critically ill patients, who have a high probability of dying soon after the emergency call is made, and for old and frail patients, who have a high probability of dying with a longer delay. Our results are in line with those found in Doyle et al. (2015) and Doyle (2011) that healthcare resources are important for emergency outcomes.

We find the largest absolute effect when evaluating 3-year mortality. The effect size is 0.4 percentage points, which (at a mean of 29%) gives an increase of 1.4%. To understand the magnitude of this effect, we use the following example. Considering that private firms service approximately 100,000 patients each year, the estimated effect suggests that private ambulance service led to an additional 400 deaths (within three years) each year. This is a considerable amount given that, for example, traffic accidents in Sweden cause a total of 200-300 deaths each year. We return to cost estimates of this in section 4.3. The next section aims to understand the mechanism behind why private ambulances underperform in quality, in both the short and long runs.

(Table 4 about here)

4.4 Mechanism – Immediate Responses and Staff Quality

We start by acknowledging the fact that mortality differences appear already after one day, even though private ambulances arrive at the scene faster than public ambulances. Although staff quality is regulated with respect to education and basic training, there are additional and other aspects of staff quality that could be important. In columns 1 and 2 in Table 5, we study the immediate actions taken by the staff when arriving at a patient's location. We know from Figure 1 that patients within a cell are comparable. Thus, any difference in assessment by the crew when arriving is likely due to crew behavior. We start by showing results using the ambulance crew's assessment of the patients' condition as the outcome. All patients approached by the crew are subjectively assigned a number from 0-7 corresponding to the severity of their condition (7 is the most severe). For this outcome, we find that private firms report that their patients are *less* severely injured than patients of public firms. This finding is particularly interesting given that we find that private firms service patients who have *higher* mortality. The possibility that private firm ambulances are matched with more severely injured patients, a bias that could explain our results on mortality, is thereby not a plausible explanation.

One explanation for this difference can be found in the next piece of evidence. An ambulance could, after examining the patient, make the decision not to bring the patient to a hospital or to any other level of care. As shown in Table 5, column 2, private firms choose more often to leave patients at home, consistent with their severity assessment. In total, 12% of all patients are non-conveyed in our data, and our estimated effect suggests that private firms leave 3 percentage points more patients at home than public firms (a 25% increase from the mean). We argue that this is strong support for immediate differences in actions taken upon arrival at the patient's location, which can explain why mortality differences appear immediately.

Leaving more patients at home could provide a possible explanation for why we find effects on mortality immediately after each dispatch. Lederman et al. (2020) studied non-conveyed patients in Stockholm in 2015 and found that among the patients who were 65 years or older, making up almost half of their adult sample, 28% visited an ED and 21% were hospitalized within 7 days of being left at home. Moreover, 0.4% of this older group died within 1 day of non-conveyance, and 1.1% died within 7 days. Private firm ambulances not conveying older patients with unclear symptoms could potentially explain at least part of the direct effects on mortality that we estimate.

Table 5 (about here)

We argue that one likely explanation for the private firm penalty in mortality comes from different recruitment strategies and consequently staff quality differences. There are arguably few other ways in which firms can cost innovate that could affect quality sufficiently much. Firms can buy cheaper vehicles and other supplies, although the ambulances must be equipped according to regulations and contracts. It is not clear to us that cheaper ambulances or supplies could have these large effects on mortality.[19] Other equipment is also tightly regulated, and ambulance station rents are normally fixed, as they

[19] One way that better vehicles could affect health is through faster transport times. What we find is that private firms reach the patients faster, and public firms reach the hospital faster. Jointly, there is not much difference in total transport times, from dispatch to drop off, between the different firms.

often share facilities with fire brigades or occupy other fixed locations. There are other more direct reasons to believe that staff quality could be important in healthcare.

Staff quality in healthcare has been shown to be very important for health outcomes. Gruber and Kleiner (2012) found that nurses' strikes in New York hospitals increased mortality among patients admitted during the strikes by 18%. This effect was similar for hospitals that engaged temporary staff, stressing the importance of location-specific experience in healthcare. Bartel et al. (2014) provide even more direct evidence on the importance of nurses' job experience and education. Departures of experienced nurses, temporarily contracted nurses, and new hires were all associated with longer care times in a US hospital environment. Emergency care (prehospital care and emergency departments, EDs) could be even more sensitive to differences in the care environment than other hospital units as a setting where both critically ill together with frail and old patients with a short life expectancy pass through. Among these patients, even small changes in the care environment can lead to large effects on health outcomes. Gruber et al. (2018) showed that setting a maximum waiting time to see a physician for patients in the ED in the UK reduced mortality by 14%.

The ambulance setting is perhaps even more complex than other in-hospital care settings. Specialist nurses (ambulance clinicians) are medically responsible for each patient and make independent decisions regarding treatments and diagnosis. Access to higher medical knowledge and second opinions are limited in the field.[20] The patients who are serviced have many kinds of ailments, diseases, and comorbidities, requiring the ambulance clinician to be a generalist with a wide range of competencies. Furthermore, ambulance staff in Stockholm receive very little feedback on their medical decisions

[20] An emergency trained physician is available on call but can be occupied with other calls, and the situation does not always allow for lengthy phone conversations.

compared to, e.g., ED nurses, who have access to diagnoses set by a physician that are based on several advanced examinations. Ambulance staff can thereby rarely evaluate whether their decisions were accurate.

Unfortunately, we have no data on the exact crew constellation for every turnout. However, we can link matched employer-employee data and calculate average quality measures of the staff for ambulance workers across private and public ambulance staff employers in Stockholm for the relevant years. Table 6 below shows the mean and the t-tests for the differences of means across provider types. As displayed rows 6 and 7, GPA and the math grade in the 9th grade indicate that public ambulances recruit higher ability staff. When studying turnover and years in the profession, the same holds true.[21] Interestingly, the salary is higher for private ambulances. However, as discussed, for example, by Besley and Ghatak (2005), it is known from the literature on compensating differentials that employment choice and wages depend on taste differences and that the public sector may well have to compensate less due to their mission-oriented organization, staffed by agents who agree with the mission. Moreover, private employees could be working more hours per year, as we do not observe hourly wages but only total income.

We interpret the evidence jointly to indicate that private firms, on average, are recruiting different competences and have difficulties retaining staff. Cost innovation at the expense of staff could cause these problems, leading these firms to be less attractive employers. There are many ways that private firms could hold costs down at the expense of employees. We have mentioned a higher rate of temporary contracts but also having more hours for a full-time contract, holding up expensive on the job training, and requiring

[21] A similar picture on tenure is found using open data from The national board of health and welfare on age among specialist ambulance nurses in Stockholm county by private/public employment status. Over most years, private firms kept more specialist nurses employed under the age of 30. See appendix Figure A1.

extra hours from their staff. Unfortunately, we have no information on these kinds of firm behaviors.

Table 6 Results: (about here)

Although the public sector seems to attract and keep more experienced staff, it is not clear how this could result in mortality gains for the patients who are serviced by public ambulances. In addition to hospital care at an ED, other alternatives are available to ambulances that could matter for patient health, especially among the fragile elderly. Direct intake at a geriatric ward is an option organized through a geriatric triage project in areas with proximity to a geriatric ward. Elderly patients who do not require emergency care at an emergency hospital could benefit from avoiding the ED. Several studies have shown that this group is particularly harmed by long waiting times in the ED caused by, e.g., patient crowding (Morley et al. 2018).

In Table 7, we study provider specific participation in health enhancing projects where patients are normally triaged and steered to an optimal level of care different from the ED. We also consider a geriatric triage project specifically where elderly could be steered directly to a geriatric ward instead of passing through the ED. 2.5% of the patients above 64 years of age were transported directly to a geriatric ward in the data. We find that private ambulances are on average less engaged in projects and that private ambulances take fewer patients directly to a geriatric ward, conditional on transporting the patients. Private ambulances do engage in more projects in their own districts compared to public ambulances in the same districts, but the average association is negative. Moreover, private ambulances are less likely to steer patients to geriatric wards when operating in both public and private districts. This suggests that it is not district specific characteristics, such as availability of geriatric wards nearby, that determines this choice, but differences in staff behaviors. Although suggestive, both due to selection of patients that are non-conveyed and also because there is missing data on project utilization, these results indicate that private ambulances are less prone to adhere to health projects, and especially to geriatric triage, which could ultimately result in poorer long-term survival of their patients.[22] Furthermore, we cannot determine if this difference is due to lower compliance to project guidelines or firm specific participation in projects for other reasons.

Table 7 (about here)

Last, we show in appendix Table A2 all other outcomes measured in seconds *after* the decision is made whether to bring the patient to the hospital or not. As the decision of whether to bring the patient to the hospital differs across provider types, the outcomes are providentially affected by selection, i.e., explained by extensive margin differences. We argue, however, with *caution,* that time at hospital still may have some informational content related to provider type. Time at hospital includes handing over the patient to the emergency department staff, resupplying the ambulance, and writing a short report. It can also include a coffee break with colleagues. We find that private firm ambulances, on average, spend 80 seconds more at hospitals and consider two main explanations for this finding since handing over the patient is normally not very time consuming. First, it could take a longer time for inexperienced staff to write the report and resupply the ambulance. Second, it could be due to private ambulance staff taking longer breaks between

[22] This analysis should be interpreted as indicative of a possible mechanism rather than a causal effect. This is because a selection of patients who are left at home has already taken place, and since much data is missing when reporting non-hospital intake and if the transport used a project such as geriatric triage.

assignments. Again, both of these explanations support the existence of differences in staff quality between firms. Moreover, if we analyze the complete transport, from dispatch call to the time the ambulance is ready for a new assignment, we do not find any significant differences. Although private ambulances reach the patient faster, the more time they spend at the other stages of the transport completely mitigates these effects. However, again, we should caution against a causal interpretation due to possible selection of patients (non-conveyance) for these outcomes.

4.5 Threats to Identification: Robustness

There are several factors that could cause threats to validity. In Figure 5, we address several issues related to ambulance services and market structure by displaying the point estimate with corresponding 95 % confidence intervals for a range of specifications. We start, again, displaying the bivariate, non-causal correlation. Using three-year mortality as the outcome in the first row, we use all the data and display the *bivariate* correlation. This correlation is negative and substantive in size, suggesting that the compositional bias between patients serviced by private firms and those serviced by the public firm is negative. This finding is expected, as private firms service a younger population, on average, as shown in Figure 2 and 4, left panel.

In the second row we show that when adding the grid fixed effects and year and month dummies, the estimated effect shifts sign and is now positive at a similar magnitude as in our preferred model. This specification could, however, be problematic and sensitive, as the allocation mechanism could be related to the priority to patients provided by the dispatch services. For ambulance dispatchers, this is the most vital piece of information they have, in addition to the location of the patient and the ambulances. Nevertheless, flexibly controlling for priority to the patient has no effect on the point estimate (row 3). This model is also the preferred model that we utilize more broadly in the paper.

Having shown how mortality responds to providers as we implement our preferred model, we now present reasonable deviations from this model to show the robustness of our results. We start by excluding the two private ambulance stations where two special ambulance units were stationed (row 3). These units were installed in 2012, have staff with higher competence, and are targeted to the most severely injured patients (nurses with anesthesiology training). While these units could spuriously relate private providers to mortality, this exclusion has only a minor impact on the estimated effect.

Since public and private ambulances operate in different areas with different characteristics, one could suspect that there are differences in the types of injuries between the areas and that this is reflected in ambulance assignment. For example, private units could, being located in areas with larger roads and faster traffic, respond to more traffic accidents and thereby face a higher probability of patient deaths. Moreover, they could be called to the scene more often as a "leading unit," organizing the care and transport of multiple injured patients. To this end, we exclude all ambulance dispatches where more than one ambulance is engaged.[23] Removing multiple unit dispatches excludes not only severe traffic accidents but also other multiple injury events, cardiac arrests, and other severe injuries where more hands or a specialist anesthesiologist nurse is required. This exclusion does not change the estimated effect (row 5).

We also exclude all priority 4 dispatches, which involve patients without a need for medical care but unable to take a taxi for different reasons. Ambulances are not supposed to transport these patients but occasionally do so. Private ambulances do so even more

[23] We do this in several ways. First, we exclude all transports where an additional unit is listed in the data. Second, we exclude all transports where there are several transports with the same case number.

often in the data. These patients, although not critically ill, have a high three-year mortality expectancy, which could feed into the overall effect of private (row 6). However, we find that this exclusion does not change the estimated relationship.

More generally, we use several approaches to challenge our estimated effect with respect to geographical or temporal confounding. First, although we have already shown that the control variables are balanced in our basic specification, we include all the control variables we have in row 7. As expected, this inclusion does not affect the estimated effect.

Another threat to identification comes from geographical changes over time. Several new ambulance units were introduced during the time period studied, and the underlying composition of residents in neighborhoods could have changed due to, e.g., housing construction. To account for secular changes over time in a flexible way, we interact the grid cell fixed effects with year-by-month dummies. In this specification, we only compare private and public ambulance dispatches within a grid cell and unique month. This restrictive specification does not change the results (row 8).

In our basic specification, we cluster at the individual level. Since treatment varies at this level, we are confident that this is the right level. However, we can also show that a conservative approach to inference, by clustering the standard errors at the grid cell level, does not render the estimated effect insignificant (row 9).

One of the most important assumptions that we make when applying our design is that the probability of being serviced by a private ambulance is close to constant within a grid cell, meaning that we have accounted for geographical selection. These rectangles are approximately 800 by 800 meters and could still contain different neighborhoods with different treatment probabilities. A grid cell located perfectly between a private and a public ambulance station could, e.g., have this characteristic. If health differs between these parts of the grid cell, our estimates could be biased. To ensure that there is limited geographical selection within the 200 by 200 grid cells, we use 1000 by 1000 grid cells and compare patients calling from a very narrow area (row 10). This exercise has no effect on the estimated effect.

Doyle et al. (2015) leverage that ambulances in the US are essentially randomly assigned to patients but use the fact that different ambulance firms have different preferences for hospitals to deliver patients to. We have argued that this is not a feature of the ambulance services in Stockholm, as patients are generally transported to a predetermined hospital based on their residence. To empirically validate this claim, we augment our preferred model with hospital fixed effects. Any hospital-specific mechanism is thereby subsumed by these indicators. Row 11 of Figure 5 shows that our claim is valid, as there is no change in the effect of introducing hospital dummies.

Moreover, as the geography of private and public ambulance stations is fixed, there could be persistent differences that affect ambulance allocation. We consider hospital transfers as such persistence. Hospitals are placed in different ambulance areas, and transferring patients between them could induce dispatch operators to selectively choose ambulances based on characteristics related to the provider. For example, transferring a patient from Huddinge Hospital (located in a private district in the southern part of the city) to an inner city hospital (there are 3 such hospitals in public districts) could imply that a public ambulance is preferred, as it will be in its home district when ready for a new assignment. By excluding transports that start at a hospital (i.e., a hospital transfer) in row 12, we find that this potential bias does not affect our estimates.

Finally, we consider time as a possible confounder for our results. If patients' health conditions are correlated with private assignment through time of the day (or day of the week), we could estimate a spurious relationship. Private ambulances could work more often in public districts at certain times or dates when patient characteristics are different.

If this is true, our estimates should be sensitive to more flexible time controls. However, including day of the week fixed effects, hour fixed effects, or even unique hour fixed effects, by considering a fixed effect for each of the 8 years by 365 days by 24 hours, does not change the results. The results remain robust to all of the challenges we are able to pose.

Figure 5: Robustness (About here)

Nevertheless, a concern could be that there are differences between private and public districts that we do not consider. The composition of patients in public areas might differ in a way that makes patient knowledge very important. If this were true, private ambulances would perform poorly in public districts but well in their own districts. To test this hypothesis, we divided Stockholm into private and public districts based on the patients' location at the time of the call. We estimate equation 1 on each of these samples using 3-year mortality as the outcome and report the results in Table 8. We find that estimating the effect of private on mortality is similar in size in both samples and that the estimates are statistically indistinguishable from each other. Furthermore, this finding is interesting because the concern that private ambulances are assigned patients in worse health must apply both in the home district and during dispatches in public districts. This would imply an assignment mechanism that is more difficult to explain.

Table 8 (about here)

Another way to make our identification strategy more credible is to focus on grid cells where the probability of being serviced by a private firm is close to the probability of being served by a public firm. If much of our identifying variation came from grid cells with poor overlap between private and public ambulances, i.e., public (private) areas where private (public) ambulances rarely go, our approach would be much more problematic. In grid cells with poor overlap, there is a higher risk of nonrandom assignment of ambulances to patients. Grid cells with an even distribution of private and public ambulances are neighborhoods with a similar distance to a public and a private ambulance station or around the borders of their respective districts. In these grid cells, the probability of being serviced by a private ambulance is close to 50%.

To approach this question, we take the average of the dummy variable *private* within each grid cell and condition our regression models on there being at least a certain fraction of private and a certain fraction of public ambulance dispatches in each cell. Figure 6 shows these results for the fraction of private dispatches in the spans of 0.1-0.9, 0.2-0.8, 0.3-0.7, and 0.4-0.6. We find that the results line up well between these different samples, even though in our most restricted sample, we only utilize approximately 10% of the full sample. These findings suggest that "border" areas provide much of the identifying variation that we use for estimation.[24]

Figure 6: Focusing on Cells Where Both Public and Private Operate (about here)

5. The Unmeasured Cost of Private Ambulances

Outsourcing public services to private firms helps the public sector to lower costs through private sector cost innovation. However, as our results show that private firms, in

[24] These areas also provide many of the observations that are used. Furthermore, since OLS implicitly weights observations on variance, cells with close to 50% private transports have a larger impact on the final estimates.

this context, produce at lower quality, cost savings from outsourcing will be lower than face value (Mukherjee, 2020).

Putting a price on mortality requires us to make several unverifiable assumptions. First, using the estimates of Table 4, mortality is increasing for three years, and then we assume that it converges between providers after 6 years. As we know that the population we study is to a large extent old and frail, this seems like a reasonable assumption. The years of life lost can thereby be calculated as,

$$Years\ Lost_t = 2 * \left(\hat{\beta}_{1Year} + \hat{\beta}_{2Years} + \hat{\beta}_{3Years}\right) * N_t^{Private}. \quad (2)$$

As private firms, each year, service approximately 100,000 patients in our data, we find that Years Lost = 2*0.01*100,000=2000. Using this number of years lost, we add a standard cost of 100,000 USD on each lost year of life (ICER, 2018). Moreover, we quality adjust each year of life lost by a constant, reflecting that those most likely to have been affected by ambulance services are old and frail. Setting this weight to 1 would imply that all years lost are years in perfect health, which is not plausible. Instead, we try different weights, 0.1 and 0.5, and calculate the total cost per year based on quality adjusted years (QALYs) of life lost.

Starting with the highest weight, we assume that each year of life lost has half the quality of years in perfect health. Multiplying the cost per year of life (100,000) by the number of years lost (2,000) and the weight (0.5) gives a total yearly cost of 100,000,000 USD, or 1000 million SEK. Using a more conservative weight gives a total yearly cost of 20,000,000 USD, or 200 million SEK, which is half the total cost of the ground-based ambulance services in Stockholm in 2012 (Årsredovisning, 2014).

Even with more conservative estimates, the total cost of years of life lost is staggering. Considering the lowest QALY weight of 0.1 and only half the lives lost (assuming that the effect is zero after three years), the yearly cost is still 10 million USD or 100 million SEK, a quarter of the total cost in 2012 (Årsredovisning, 2014). Even if we take a cautious and conservative approach, our calculations suggest that the cost of reduced quality vastly outweighs the cost savings that private firms generate.

6 Conclusion

Outsourcing public services to private firms is a tool to help the public sector contain costs without reducing quality. Theoretically, the economics literature describes a trade-off related to the contractibility of quality in public domains considered for outsourcing. This trade-off comes from private firms' strong incentives to cost innovate but weak incentives to uphold quality if it is not explicitly profitable. When private firms are not sufficiently incentivized in contracts to produce a minimum level of quality, it is possible, or even likely, that they will produce lower quality. When quality can be contracted on, private firms can theoretically produce adequate quality at a lower cost than public sector in-house production.

We test these predictions in the health care sector, a sector that has previously received little attention. Health outcomes are often difficult to contract on, as health outcomes may reflect the composition of patients rather than provider quality. We solve this challenge in an ambulance services environment where both private and public firms operate side-by-side. We confirm the theoretical predictions that private firms can uphold quality if contracted on but produce substantially lower quality on non-contracted health outcomes. Patients serviced by private ambulances have higher mortality already within one day and up to three years after being serviced.

We provide evidence in favor of the explanation that the difference in staff quality is a likely reason for the mortality differences. Moreover, additional evidence shows behavioral differences between ambulance firms during each assignment. In particular, we find that private firm ambulances leave substantially more patients at home than public ambulances. Causal estimates are complemented by descriptive evidence suggesting that private firms have a higher turnover of staff and that they keep younger and less experienced staff. Interestingly, private employees have higher annual incomes, which could reflect both a wage premium to work for these firms or the possibility that private employees are working more hours. Nevertheless, private firms' cost innovation at the expense of their staff likely affects quality of care.

When quality is difficult to contract on, theory advice against outsourcing. Policy makers prone to outsource try to solve this dilemma by contracting on proxies or on inputs for quality. However, this procedure is subject to moral hazard and provides strong incentives for firms to comply with the contracted measures while shirking on unmeasured aspects of quality. Moreover, it is worth noting that writing too specific contracts on quality proxies may not be the solution as it will severely limit private firms' ability to cost innovate through competition. Contracting on the outcome is thereby the best option to uphold quality and take advantage of private firms' inherent superiority in cost innovation. If that is not possible, theory and our evidence show that outsourcing may have very adverse effects. Our estimated effect suggests that private ambulance service led to an additional 400 deaths (within three years) each year. Our back-of-the-envelope cost calculations suggest that the loss due to private ambulance firms is substantial. Even a conservative estimate finds that the cost could be approximately one quarter to a half of the total cost of the ambulance services together. Thus, we are unconvinced that outsourcing is a superior option in this and similar contexts and our policy recommendation would be to abstain from outsourcing in the ambulance sector.

References

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. (2017). “When should you adjust standard errors for clustering?” Working Paper 24003, *National Bureau of Economic Research*.

AISAB (2005). Förfrågningsunderlag upphandling av Ambulanssjukvårdstjänst. Stockholm Läns Landsting.

AISAB (2020). Ambulanssjukvården i Stockholm AB. https://www.aisab.nu/om-oss/nyheter/aisab-20-ar/ Webpage accessed: 20200217.

Andersson, F., Jordahl, H., and Josephson, J. (2019). “Outsourcing Public Services: Contractibility, Cost, and Quality.” CESifo Economic Studies. ifz009.

Armstrong, G. S. and D. L. MacKenzie (2003), “Private versus Public Juvenile Correctional Facilities: Do Differences in Environmental Quality Exist?” *NCCD News* 49, 542–63.

Athey, S. and Stern, S. (2002). “The impact of information technology on emergency health care outcomes.” The RAND Journal of Economics, 33(3):399–432.

Barron, D. N. and West, E. (2017). “The quasi-market for adult residential care in the UK: Do for-profit, not-for-profit or public sector residential care and nursing homes provide better quality care?” *Social Science & Medicine*, 179:137 – 146.

Bartel, A. P., Beaulieu, N. D., Phibbs, C. S., and Stone, P. W. (2014). “Human capital and productivity in a team environment: Evidence from the healthcare sector.” American Economic Journal: Applied Economics, 6(2):231–59.

Bayer, P. and D. Pozen (2005), “The Effectiveness of Juvenile Correctional Facilities: Public versus Private Management”, *Journal of Law and Economics* 48, 549–89.

Behaghel, L., Crepon, B., and Gurgand, M. (2014). “Private and public provision of counseling to job seekers: Evidence from a large controlled experiment.” *American Economic Journal: Applied Economics*, 6(4):142–174.

Bennmarker, H., Grönqvist, E., and Öckert, B. (2013). ”Effects of contracting out employment services: Evidence from a randomized experiment.” *Journal of Public Economics*, 98:68 – 84.

Bergman, M. A., Johansson, P., Lundberg, S., and Spagnolo, G. (2016). “Privatization and quality: Evidence from elderly care in Sweden.” *Journal of Health Economics*, 49:109 –119.

Besley, Timothy, and Maitreesh Ghatak. 2005. "Competition and Incentives with Motivated Agents." *American Economic Review*, 95 (3): 616-636.

DeWenter, Kathryn, L., and Paul H. Malatesta. 2001. "State-Owned and Privately Owned Firms: An Empirical Analysis of Profitability, Leverage, and Labor Intensity." *American Economic Review*, 91 (1): 320-334.

Devereaux, P., Choi, P. T., Lacchetti, C., Weaver, B., Schünemann, H. J., Haines, T., Lavis, J. N., Grant, B. J., Haslam, D. R., Bhandari, M., Sullivan, T., Cook, D. J., Walter, S. D., Meade, M., Khan, H., Bhatnagar, N., and Guyatt, G. H. (2002). “A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals”. Canadian Medical Association Journal, 166(11):1399–1406.

Doyle, J. J. (2011). “Returns to local-area health care spending: Evidence from health shocks to patients far from home.” American Economic Journal: Applied Economics, 3(3):221–43.

Doyle, J. J., Graves, J. A., and Gruber, J. (2017). “Evaluating measures of hospital quality.” Working Paper 23166, National Bureau of Economic Research.

Doyle, J. J., Graves, J. A., Gruber, J., and Kleiner, S. A. (2015). “Measuring returns to hospital care: Evidence from ambulance referral patterns.” *Journal of Political Economy*, 123(1):170–214.

Duggan, Mark, Jonathan Gruber, and Boris Vabson. 2018. "The Consequences of Health Care Privatization: Evidence from Medicare Advantage Exits." *American Economic Journal: Economic Policy*, 10 (1): 153-86.

Duwe, G. and V. Clark (2013), “The Effects of Private Prison Confinement on Offender Recidivism: Evidence from Minnesota”, *Criminal Justice Review* 38, 375–94

Farabee, D. and K. Knight (2002), “A Comparison of Public and Private Prisons in Florida: During- and Post-Prison Performance Indicators”, *Query Research*, Los Angeles, CA.

Grabowski, D. C., Feng, Z., Hirth, R., Rahman, M., and Mor, V. (2013). “Effect of nursing home ownership on the quality of post-acute care: An instrumental variables approach.” Journal of Health Economics, 32(1):12 – 21.

Gruber, J., Hoe, T. P., and Stoye, G. (2018). “Saving lives by tying hands: The unexpected effects of constraining health care providers.” Working Paper 24445, National Bureau of Economic Research.

Gruber, J. and Kleiner, S. A. (2012). “Do strikes kill? Evidence from New York state.” American Economic Journal: Economic Policy, 4(1):127–57.

Hart, Oliver, and John Moore. 1990. “Property rights and the nature of the firm.” *Journal of Political Economy*. 98(6): 1119-1158.

Hart, Oliver, Shleifer, Andrei and Vishny, Robert W, “The Proper Scope of Government: Theory and an Application to Prisons”, *The Quarterly Journal of Economics*, Volume 112, Issue 4, November 1997, Pages 1127–1161.

Hälso och sjukvårdsförvaltningen (2011). ”Upphandling av ambulanssjukvården. faktasammanställning från hälso- och sjukvårdsnämndens förvaltning.” *Technical report, Stockholms läns landsting.*

Hälso och sjukvårdsförvaltningen (2018). ”Årsrapport 2017. Prehospitala verksamheter i SLL.” *Technical Report 2018:0024,* Stockholms läns landsting.

Hälso och sjukvårdsförvaltningen (2014). ” Årsrapport 2014 prehospitala verksamheter i SLL.” Technical Report 2014, Stockholms läns landsting.

Hull, P. (2018). ”Estimating hospital quality with quasi-experimental data.” Mimeo.

Jansson, A. and Castwall, H. (2017). ”Är privat ambulansverksamhet effektivare än offentlig - En analys av de prehospitala ambulansvårdgivarna i Stockholms län under år 2015.” Kandidatuppsats i nationalekonomi. HT2016.

Jena, Anupam B., N. Clay Mann, Leia N. Wedlund, and Andrew Olenski. (2017) "Delays in emergency care and mortality during major US marathons." New England Journal of Medicine 376, no. 15. 1441-1450.

Krueger, A. B. (1999). ”Experimental Estimates of Education Production Functions.” *The* Quarterly Journal of Economics, 114(2):497–532.

Landstingsrevisorerna (2013). “Prehospital vård – vårdkedjans första insats”. Stockholms Läns landsting Projektrapport Nr. 08/2013

Laun, L. and P. Skogman Thoursie (2014), “Does Privatisation of Vocational Rehabilitation Improve Labour Market Opportunities? Evidence from a Field Experiment in Sweden”, *Journal of Health Economics* 34, 59–72.

Lee, Michael H., Jeremiah D. Schuur, and Brian J. Zink. "Owning the cost of emergency medicine: beyond 2%." Annals of emergency medicine, 62.5 (2013): 498-505.

Lucchese, Elena. (2020) “It Could be Worse...It Could be Raining: Ambulance Response Time and Health Outcomes” University of Milan Bicocca Department of Economics, Management and Statistics Working Paper No. 429.

Mukherjee, Anita. 2020. "Impacts of Private Prison Contracting on Inmate Time Served and Recidivism", *American Economic Journal: Economic Policy*.

Petersen, O. H., Hjelmar, U., and Vrangbæk, K. (2018). “Is contracting out of public services still the great panacea? a systematic review of studies on economic and quality effects from 2000 to 2014.” *Social Policy & Administration*, 52(1):130–157.

Powers, R., C. Kaukinen, and M. Jeanis (2017), “An Examination of Recidivism among Inmates.Released from a Private Reentry Center and Public Institutions in Colorado”, The Prison Journal 97, 609–27.

Riksrevisionen (2012) ”Statens insatser inom Ambulansverksamheten” . *RiR* 2012:20

Shleifer, A., 1998. “State versus private ownership.” *Journal of Economic Perspectives*, 12(4), 133–150.

SLL (2017). Årsprotokoll från uppföljning 2017 – AISAB/Falck/Samariten, vägburen ambulans. Hälso och sjukvårdsförvaltningen, Stockholms läns Landsting.

SOS Alarm (2020). https://www.sosalarm.se/Kund/tjanstekatalog2/ambulansdirigering/

Spivak, A. L. and S. F. Sharp (2008), “Inmate Recidivism as a Measure of Private Prison Performance”, Crime and Delinquency 54, 482–508.

Stolt, R., Blomqvist, P., and Winblad, U. (2011). ”Privatization of social services: Quality differences in Swedish elderly care.” Social Science & Medicine, 72(4):560 – 567.

Szymanski, S. (1996), “The Impact of Compulsory Competitive Tendering on Refuse Collection Services”, *Fiscal Studies* 17, 1–19.

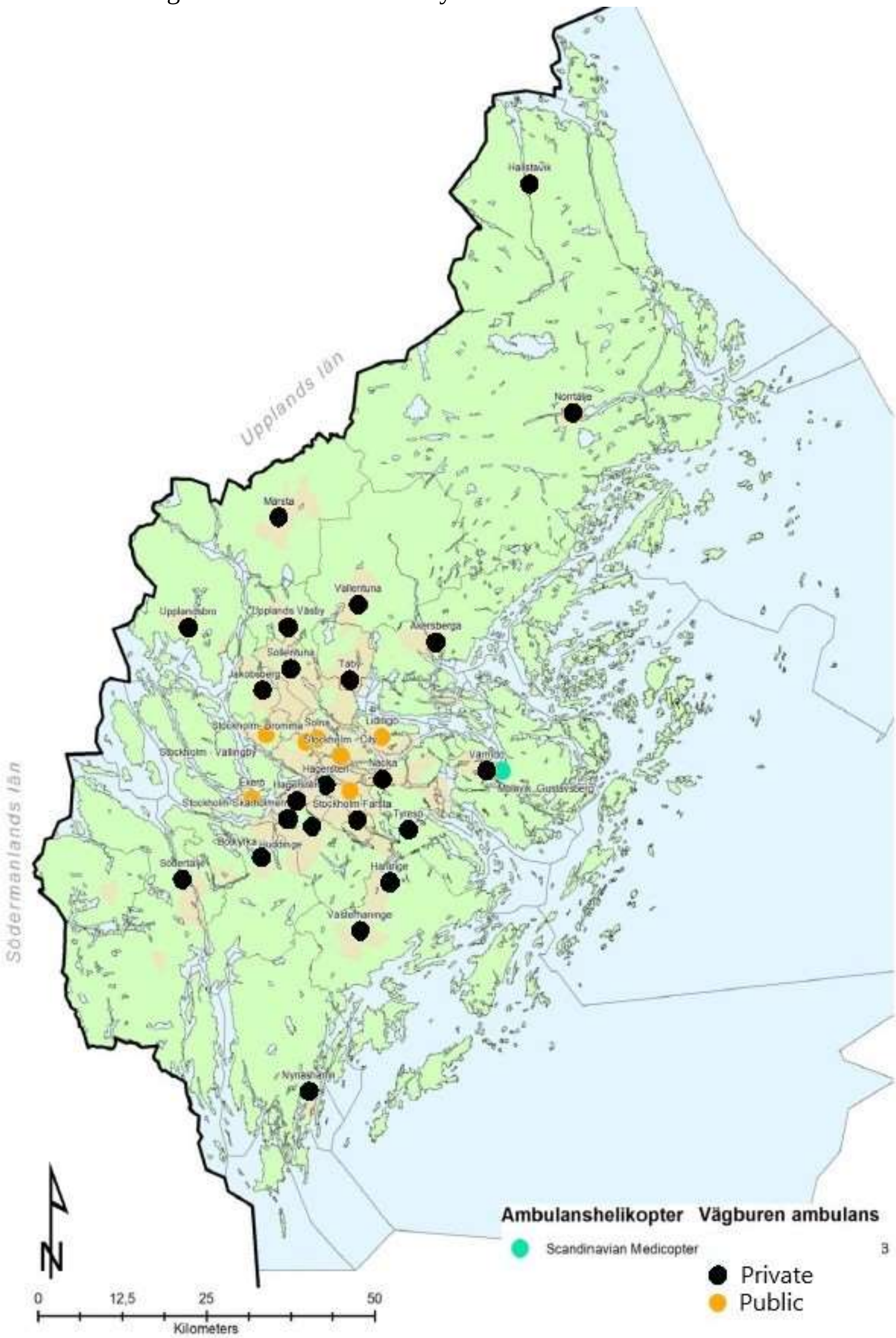
Szymanski, S. and S. Wilkins (1993), “Cheap Rubbish? Competitive Tendering and Contracting out in Refuse Collection – 1981–88*”, Fiscal Studies* 14, 109–30.

Tanuseputro, P., Chalifoux, M., Bennett, C., Gruneir, A., Bronskill, S. E., Walker, P., and Manuel, D. (2015). “Hospitalization and mortality rates in long-term care facilities: Does for-profit status matter?” *Journal of the American Medical Directors Association*, 16(10):874 – 883.

Wübker, Ansgar and Wuckel, Christiane. “The Impact of Private for-Profit Hospital Ownership on Costs and Quality of Care – Evidence from Germany”, *CESifo Economic Studies*, Volume 65, Issue 4, December 2019, Pages 373–401.

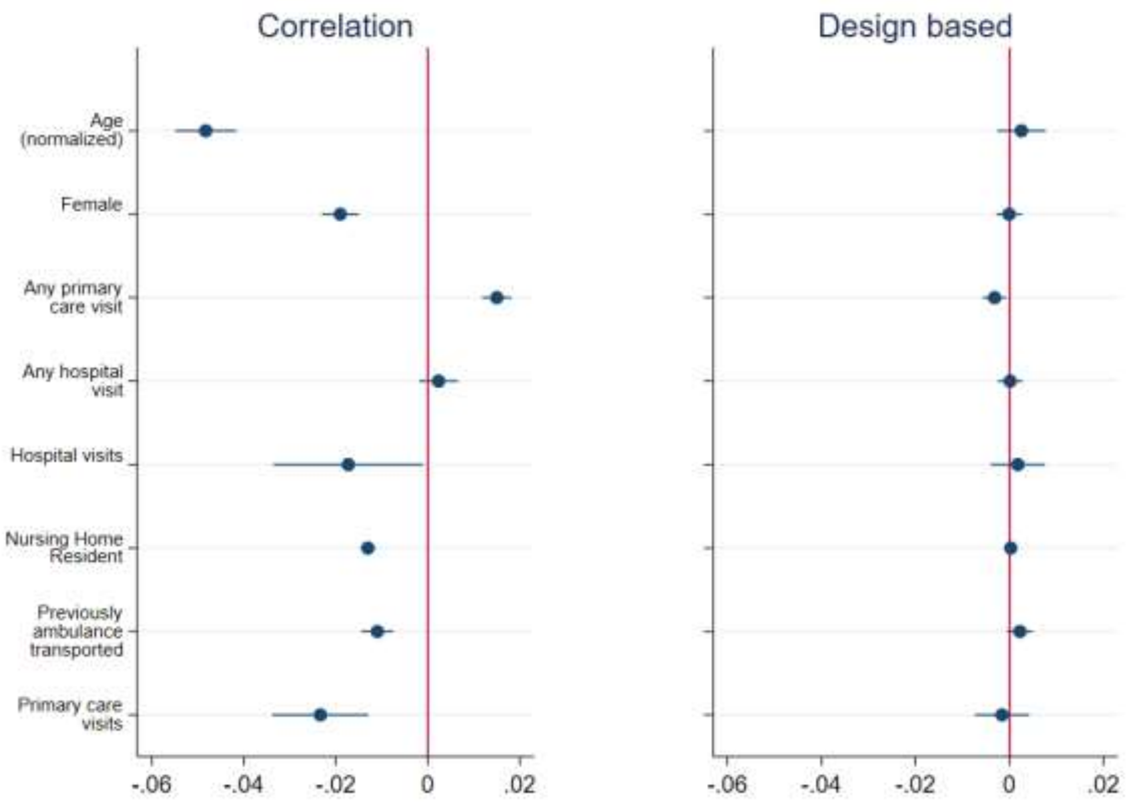
Figures and Tables

Figure 1: Stockholm County with Ambulance Stations



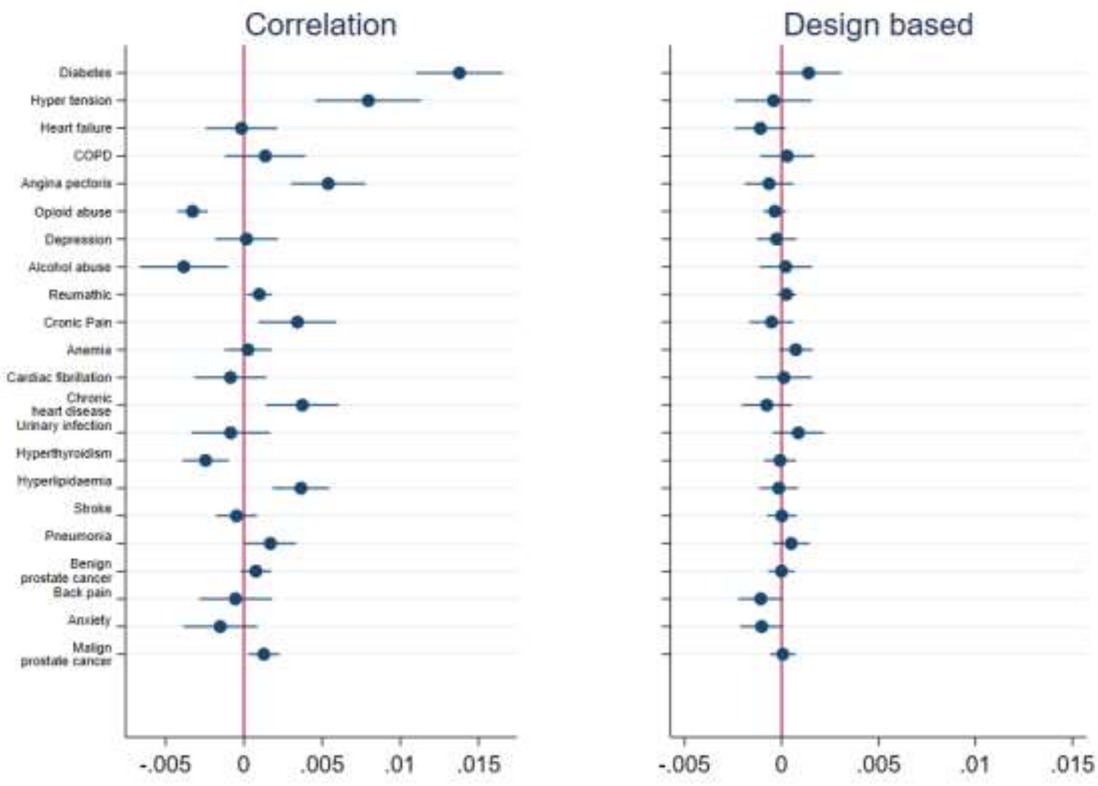
Note: Location of ambulance stations in Stockholm county by private (black dots) and public (orange dots) providers. Reproduced from Hälso och sjukvårdsförvaltningen (2014).

Figure 2. Balance of predetermined variables: Bivariate and design based



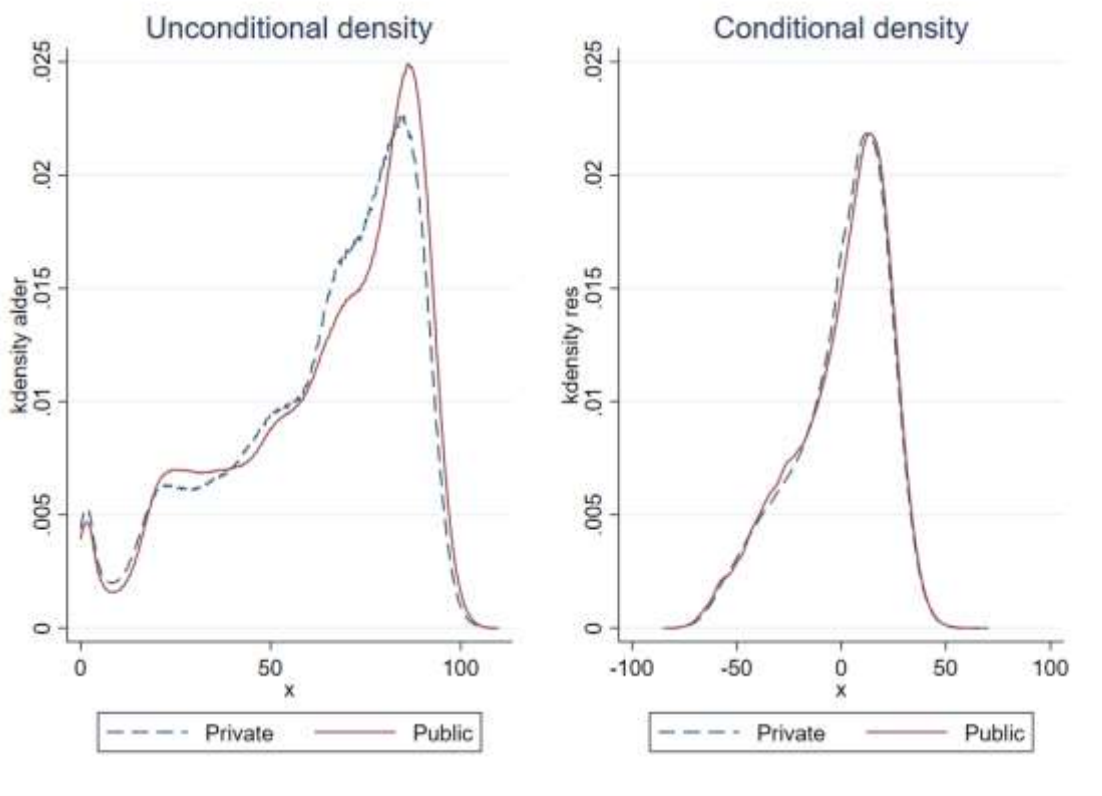
Note: Both panels show estimates of private with the predetermined variable indicated to the left of the figures as the outcome. The left panel shows the bivariate linear regression estimate using OLS (only the outcome and private indicator), while the right panel displays the same results using our preferred model. Several variables here are measured at the time of each ambulance dispatch. Age in years (normalized by demeaning and dividing by the standard deviation), sex, residence in a nursing home, and previous transport by an ambulance (after 2009) have this property. Any visit and number of visits to primary care (hospital care) is based on data from 2007-2008.

Figure 3: Balance of predetermined variables: Bivariate and design based



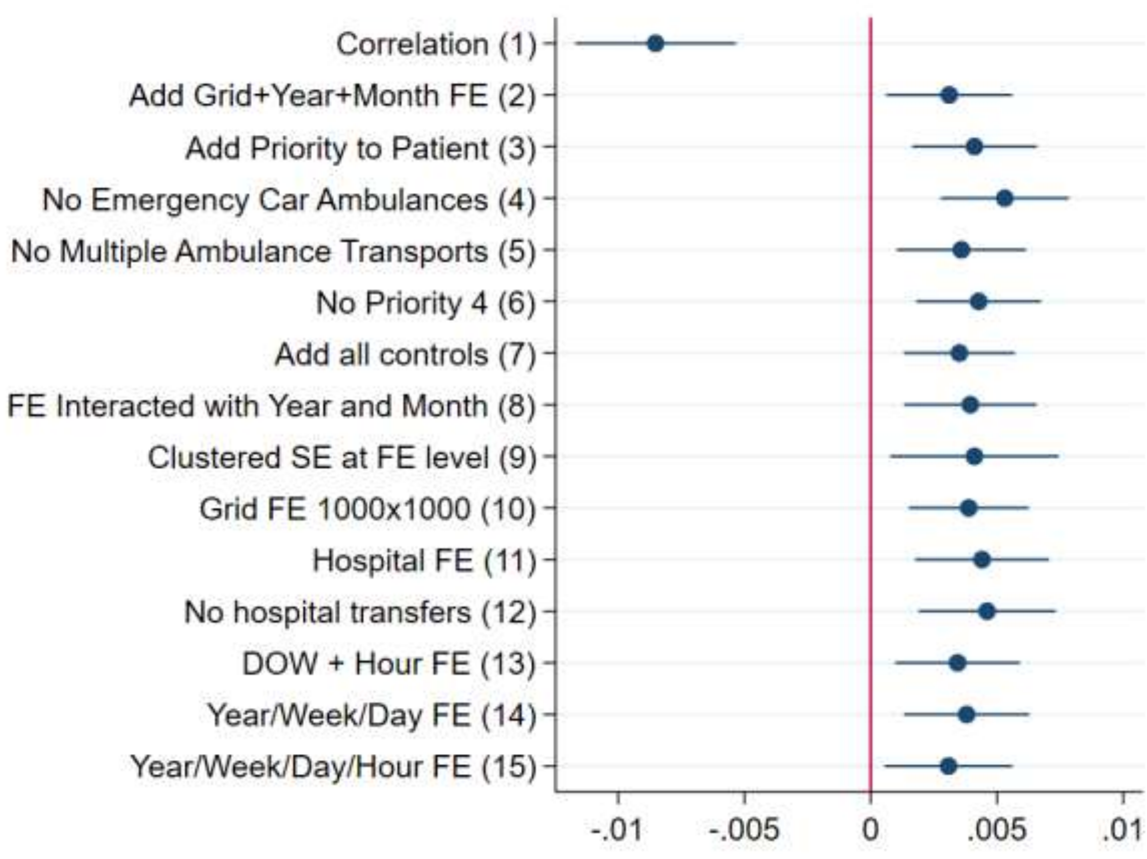
Note: Both panels show estimates of private with the predetermined variable indicated to the left of the figures as the outcome. The left panel shows the bivariate linear regression estimate using OLS (only the outcome and the private indicator), while the right panel displays the same results using our preferred model. Diseases are based on ICD codes registered at hospital and primary care visits in 2007-2008. If a diagnosis is documented, the patient indicator variable for that disease is 1. If there is no indication, it is 0.

Figure 4: Differences in the Age Distribution of Patients



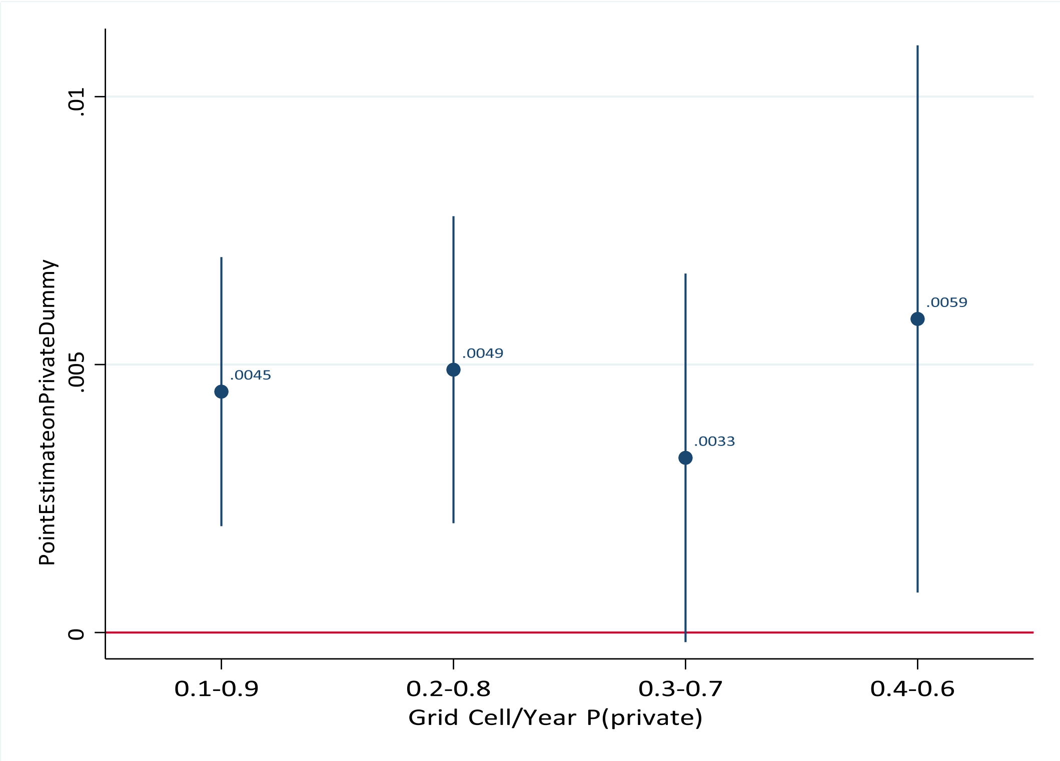
Note: Kernel density plots of patient age distribution (left panel) and conditional age distribution (right panel), between private and public firms. The conditional distribution is residualized by our preferred specification using 200x200 grid fixed effects.

Figure 5 Robustness



Note: This figure presents point estimates and 95% confidence intervals for different specifications of equation 1. The first three rows impose fewer restrictions than equation 1. The first row is the raw correlation between 3-year mortality and private ambulance providers. Row 2 add fixed effects, leading to our preferred specification (row 3). Rows 4 and above challenge this specification by adding structure or restricting the sample in different ways, as described in the text.

Figure 6: Focusing on Cells Where Both Public and Private Firms Operate



Note: P(private) is generated by taking the mean of *Private* at the grid cell (1000 x 1000) and year level. The regressions described above are baseline but restricted to year cells with public/private overlap as defined above. For example, $0.1 - 0.9$ means that only year/cells with at least 10 percent private (public) are included in the regression. The number of observations ranges from around 1 million (on the left) to 100 k (on the right).

Table 1: Summary Statistics 1: Health

	Mean	Std.	Min	Max	Obs
Predetermined characteristics					
Previously serviced by ambulance <365 days	0.453	0.498	0.00	1.00	1141939
Age	62.251	24.844	0.00	110.00	1138319
Female	0.533	0.499	0.00	1.00	1141855
Nursing home resident	0.070	0.254	0.00	1.00	1141939
Diagnoses					
Any health center visit	0.70	0.46	0.00	1.00	1141939
Any hospital visit	0.38	0.48	0.00	1.00	1141939
Health center visits	37.26	62.57	0.00	400.00	1141939
Hospital visits	1.44	3.95	0.00	114.00	1141939
Diabetes	0.10	0.30	0.00	1.00	1141939
Hypertension	0.15	0.36	0.00	1.00	1141939
Heart failure	0.06	0.23	0.00	1.00	1141939
COPD	0.07	0.25	0.00	1.00	1141939
Angina pectoris	0.05	0.22	0.00	1.00	1141939
Opioid disease	0.01	0.09	0.00	1.00	1141939
Depressive disease	0.03	0.18	0.00	1.00	1141939
Alcohol disease	0.06	0.23	0.00	1.00	1141939
Rheumatic	0.01	0.09	0.00	1.00	1141939
General pain	0.04	0.20	0.00	1.00	1141939
Anemia	0.02	0.14	0.00	1.00	1141939
Cardiac fibrillation	0.07	0.25	0.00	1.00	1141939
Chronic heart disease	0.05	0.22	0.00	1.00	1141939
Urinary infection	0.06	0.23	0.00	1.00	1141939
Hyperthyroidism	0.02	0.15	0.00	1.00	1141939
Hyperlipidemia	0.03	0.18	0.00	1.00	1141939
Stroke	0.02	0.14	0.00	1.00	1141939
Pneumonia	0.03	0.17	0.00	1.00	1141939
Benign prostate cancer	0.02	0.12	0.00	1.00	1141939
Back pain	0.04	0.20	0.00	1.00	1141939
Anxiety	0.04	0.19	0.00	1.00	1141939
Malign prostate cancer	0.01	0.12	0.00	1.00	1141939

Note: Data from the Stockholm county VAL database. The “Previously transported by ambulance” variable is restricted to a year before the index dispatch.

Table 2: Summary Statistics 2: Outcomes

	Mean	Std.	Min	Max	Obs
Treatment					
Private ambulance	0.587	0.492	0.00	1.00	1141939
Private ambulance Res	-0.000	0.342	-1.05	1.01	1088143
Priority to patient	1.715	0.699	1.00	4.00	1141820
Outcomes: Ambulance responses					
Patient queue time (s)	3134.43	8717.76	0.00	36000.00	1141939
Response time (s)	106.82	224.00	0.00	1800.00	1059247
Time to patient (s)	779.34	525.68	0.00	3600.00	1073220
Time at patient (s)	1083.22	685.90	0.00	14335.00	1013065
Time to hospital (s)	1014.47	634.85	0.00	7200.00	969998
Time at hospital (s)	1885.44	790.91	0.00	7196.00	884408
Assessed severity	2.84	0.99	0.00	7.00	1125465
Priority to hospital	2.16	0.68	1.00	4.00	1003289
Patient to hospital	0.88	0.33	0.00	1.00	1141939
Outcomes: Health					
Mortality 3 years	0.29	0.45	0.00	1.00	1141939
Mortality 3y Res	-0.00	0.43	-1.17	1.09	1088143
Mortality 2 years	0.24	0.43	0.00	1.00	1141939
Mortality 1 year	0.18	0.38	0.00	1.00	1141939
Mortality 6 months	0.13	0.34	0.00	1.00	1141939
Mortality 3 months	0.10	0.30	0.00	1.00	1141939
Mortality 1 month	0.06	0.24	0.00	1.00	1141939
Mortality 1 week	0.03	0.18	0.00	1.00	1141939
Mortality 1 day	0.02	0.13	0.00	1.00	1141939

Source: Data from the Stockholm county VAL database and SOSAlarm AB. Treatment variable and main outcome (postfix “Res” indicates residualized variables using our main specification with fixed effects.)

Table 3: Contracted outcomes

	Response time to a dispatch	Travel time to patient
	(1)	(2)
Private ambulance	-8.3980	-63.7659
	(0.7620)	(1.4493)
Public outcome mean	107.6759	716.5211
Observations	1012999	1029854

Note: Each model includes 200x200 grid fixed effects covering the entire county of Stockholm, where each grid is approximately 700x600 meters large, as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excludes extra ambulances. Standard errors are clustered at each patient.

Table 4: Noncontracted outcomes

	Mortality after ambulance dispatch occurred within:							
	1 day	1 week	1 month	3 months	6 months	1 year	2 years	3 years
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Private ambulance	0.0011	0.0009	0.0013	0.0021	0.0009	0.0022	0.0037	0.0041
	(0.0003)	(0.0005)	(0.0007)	(0.0008)	(0.0009)	(0.0011)	(0.0012)	(0.0013)
Public outcome mean	0.0148	0.0294	0.0581	0.0954	0.1290	0.1758	0.2449	0.2935
Observations	1088143	1088143	1088143	1088143	1088143	1088143	1088143	1088143

Note: Each model includes 200x200 grid fixed effects covering the entire county of Stockholm, where each grid is approximately 700x600 meters large, as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excludes extra ambulances. Standard errors are clustered at each patient.

Table 5: Does firm affiliation affect ambulance crew behavior?

	Estimated severity	Patient to hospital
	(1)	(2)
Private ambulance	-0.0534	-0.0328
	(0.0027)	(0.0009)
Public outcome mean	2.8507	0.8990
Observations	1073846	1088143

Note: Each model includes 200x200 grid fixed effects covering the entire county of Stockholm, where each grid is approximately 700x600 meters large, as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excludes extra ambulances. Standard errors are clustered at each patient.

Table 6: Firms and Quality

	Observations	Private Mean	Public Mean	Difference	t-value
Graduation Year	10	2005.904	2005.669	.2350464	.4791087
Experience	10	6.573536	6.831337	-.2578012	-.5300849
Turnover	9	.1802916	.0936577	.0866339	1.859895
Years in Profession	10	4.629704	5.020975	-.3912717	-2.323079
Age	10	39.50749	40.05018	-.5426849	-1.085429
9th Grade GPA	10	.0705313	.412161	-.3416297	-4.580257
9th Grade Math	8	.2425694	.6664265	-.4238571	-3.469922
Wage Income	10	430381.4	394543.9	35837.47	3.429978

Note: Each row describes a t-test between public and private ambulance firms in Stockholm. The variables are aggregated to the yearly level before the test and each year measure 100-200 individuals working in the firms.

Table 7: Is provider related to utilization of health enhancing projects?

	All projects			Geriatric project		
	All valid data	Private area	Public area	Above age 64	Private area	Public a.
	(1)	(2)	(3)	(4)	(5)	(6)
Private ambulance	-0.0012	0.0016	-0.0051	-0.0050	-0.0034	-0.0074
	(0.0005)	(0.0007)	(0.0007)	(0.0006)	(0.0008)	(0.0008)
Public outcome mean	0.0266	0.0264	0.0266	0.0247	0.0221	0.0256
Observations	798627	489379	309104	449976	276057	173799

Note: This information comes from the ambulance data where a field called, "Projekt" should be filled in if a project is utilized during the transport. A project is a local or global trial where new methods are tried and evaluated. Columns 2 and 5 restrict the sample to private districts only while columns 3 and 6 only use patients serviced in public ambulance districts. Columns 4-6 further restricts the sample to patients above age 64. Each model include 200x200 grid fixed effects covering the entire county of Stockholm where each grid is approximately 700x600 meters large as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excluding extra ambulances. Standard errors are clustered at each patient.

Table 8: Does the effect depend on the private or public area of the patient?

	Public area	Private area
	(1)	(2)
Private Ambulance	0.0033	0.0052
	(0.0019)	(0.0016)
Public outcome mean	0.2952	0.2883
Observations	419937	668014

Note: Each model includes 200x200 grid fixed effects covering the entire county of Stockholm, where each grid is approximately 700x600 meters large, as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excludes extra ambulances. Standard errors are clustered at each patient.

Appendix

Table A1: Balance of predetermined variables

	Variable mean	Estimate of Private on variable	
		Private - Public Bivariate	Private - Public Basic model
	(1)	(2)	(3)
Previously transported by ambulance <1 year	0.4530	-0.0110 (0.0018)	0.0022 (0.0014)
Age	62.2510	-1.1992 (0.0844)	0.0628 (0.0656)
Female	0.5334	-0.0190 (0.0020)	-0.0001 (0.0014)
Nursing home resident	0.0695	-0.0130 (0.0008)	0.0003 (0.0007)
Any health center visit	0.6998	0.0150 (0.0016)	-0.0032 (0.0013)
Any hospital visit	0.3775	0.0023 (0.0022)	0.0001 (0.0014)
Health center visits	37.2555	-1.4624 (0.3356)	-0.1001 (0.1855)
Hospital visits	1.4369	-0.0683 (0.0328)	0.0070 (0.0116)
Diabetes	0.1009	0.0138 (0.0014)	0.0014 (0.0009)
Hypertension	0.1499	0.0080 (0.0017)	-0.0004 (0.0010)
Heart failure	0.0578	-0.0002 (0.0012)	-0.0011 (0.0007)
COPD	0.0652	0.0014 (0.0013)	0.0003 (0.0007)
Angina pectoris	0.0527	0.0054 (0.0012)	-0.0006 (0.0006)
Opioid disease	0.0082	-0.0033 (0.0005)	-0.0004 (0.0003)
Depressive disease	0.0336	0.0002 (0.0010)	-0.0003 (0.0005)
Alcohol disease	0.0560	-0.0038 (0.0014)	0.0002 (0.0007)
Rheumatic	0.0089	0.0010 (0.0004)	0.0002 (0.0003)
Chronic pain	0.0404	0.0034 (0.0013)	-0.0005 (0.0006)
Anemia	0.0211	0.0003 (0.0008)	0.0007 (0.0004)
Cardiac fibrillation	0.0696	-0.0009 (0.0012)	0.0001 (0.0007)
Chronic heart disease	0.0530	0.0037 (0.0012)	-0.0008 (0.0007)
Urinary infection	0.0579	-0.0008 (0.0013)	0.0009 (0.0007)
Hyperthyroidism	0.0223	-0.0024 (0.0008)	-0.0001 (0.0004)
Hyperlipidemia	0.0327	0.0037 (0.0009)	-0.0001 (0.0005)
Stroke	0.0203	-0.0005 (0.0007)	0.0000 (0.0004)
Pneumonia	0.0291	0.0017 (0.0008)	0.0005 (0.0005)
Benign prostate cancer	0.0156	0.0008 (0.0005)	-0.0000 (0.0003)
Back pain	0.0435	-0.0005 (0.0012)	-0.0011 (0.0006)
Anxiety	0.0361	-0.0015 (0.0012)	-0.0010 (0.0006)
Malign prostate cancer	0.0149	0.0013 (0.0005)	0.0001 (0.0003)

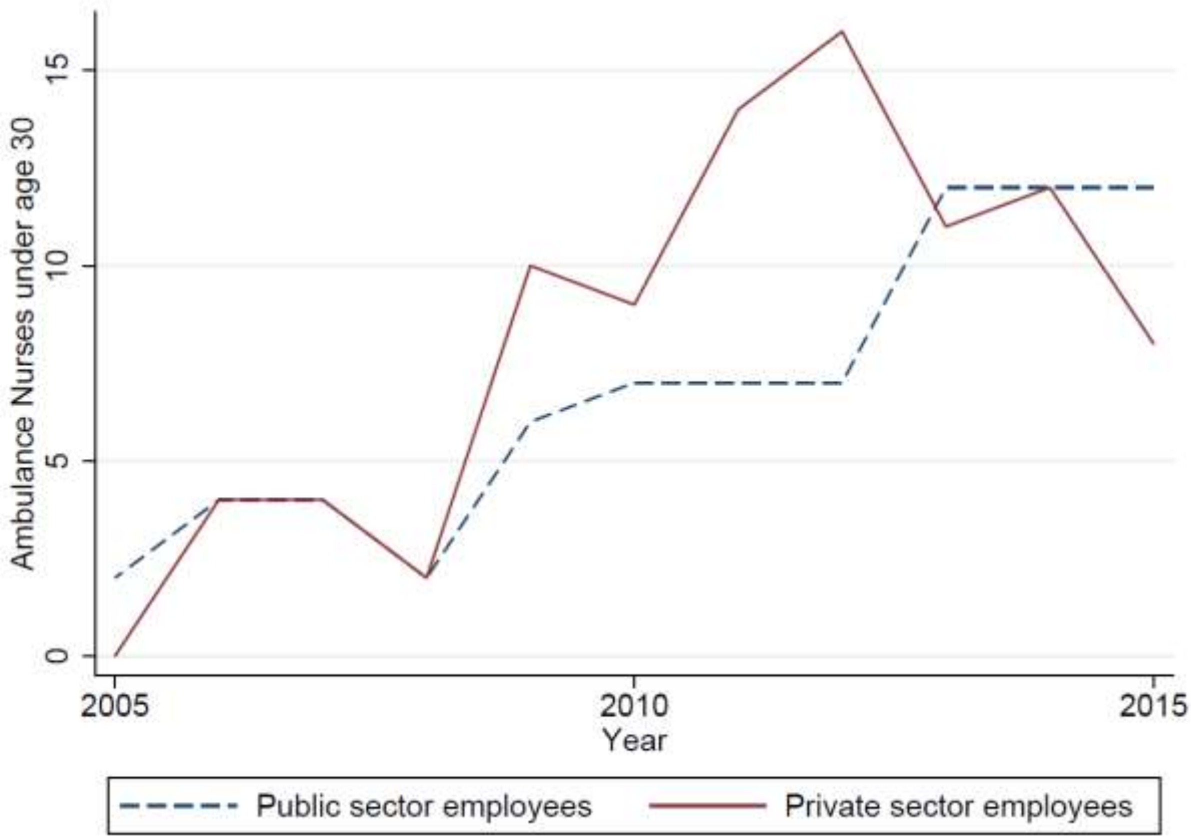
Note: Each model includes 200x200 grid fixed effects covering the entire county of Stockholm, where each grid is approximately 700x600 meters large, as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excludes extra ambulances. Standard errors are clustered at each patient.

Table A2: Using all time stamps as outcomes

	Response time	Time to patient	Time at patient	Time to hospital	Time at hospital
	(1)	(2)	(3)	(4)	(5)
Private ambulance	-8.3980	-63.7659	44.8114	46.3217	86.4880
	(0.7620)	(1.4493)	(1.9917)	(1.3591)	(2.4349)
Public outcome mean	107.6759	716.5211	1079.8855	878.0634	1860.1842
Observations	1012999	1029854	964109	924965	846380

Note: Each model include 200x200 grid fixed effects covering the entire county of Stockholm where each grid is approximately 700x600 meters large as well as year and month dummies and priority to patient dummies. The sample is restricted to ground emergency ambulances and excluding extra ambulances. Standard errors are clustered at each patient.

Figure A1. Number of young nurses employed by sector



Note: Number of trained nurses with specialization in pre-hospital care under age 30 employed by public and private firms in Stockholm county 2005-2015. Source: Open data from National board of health and welfare (Socialstyrelsen).